La mente y el cosmos

Por qué la concepción neo-darwinista materialista de la naturaleza es, casi con certeza, falsa

Thomas Nagel



Tomas Nagel defiende lo insostenible de las versiones hoy día más usuales del naturalismo y afirma que la versión materialista estándar de la biología evolutiva es, sin remedio, incompleta. La historia que condujo al origen de la vida y a la evolución de las especies no ha podido ser un proceso regido exclusivamente por los principios de la selección natural darwiniana…



Thomas Nagel

La mente y el cosmos

Por qué la concepción neo-darwinista meterialista de la naturaleza es, casi con certeza, falsa

ePub r1.0
Titivillus 01.09.17

Título original: *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*
Thomas Nagel, 2012
Traducción: Francisco Rodríguez Valls

Editor digital: Titivillus
ePub base r1.2

PRÓLOGO

Es extraño que un libro tan heterodoxo como el que el lector tiene entre sus manos llegue al público en español tan solo un par de años después de haberse publicado en su lengua original. No es precisamente la última de las razones de por qué eso no ocurre con más frecuencia el hecho de que nuestro panorama intelectual esté fuertemente sesgado hacia lo políticamente correcto y en él se dé prioridad a otras obras que abundan más en los motivos, a veces excesivamente rebuscados y contrarios al sentido común, defendidos por el naturalismo reduccionista en ciencia natural. De esta manera se hurtan al debate en castellano otros enfoques de envergadura. La minoría que defiende esos enfoques diferentes tiene mucho que decir para transformar una ciencia demasiado estancada en los últimos decenios en ciertas áreas, aunque en nuestro ámbito esas publicaciones intenten ocultarse tanto al estamento intelectual como al público en general.

El propósito del libro de Nagel, posiblemente uno de los mejores y más influyentes filósofos vivos en el panorama contemporáneo de la filosofía de la mente, es una crítica a aquella expresión que acuñó un autor clásico inglés del siglo XIX, de cuyo nombre no quiero acordarme, y que se resume en las siguientes ideas: la ciencia actual dice que puede explicarlo todo y, cuando ve que hay algo que no puede explicar, dice que ese algo o no existe o es un accidente irrelevante. El libro de Nagel no es, como algunos equivocadamente han afirmado, antinaturalista sino, ante todo, antirreduccionista y para serlo aduce unas razones sumamente convincentes que tienen en cuenta, con el propósito de superarlos, los límites que se autoimpuso la ciencia natural en la revolución científica y que, sin duda, la han llevado muy lejos. Pero los árboles de esos límites le están impidiendo ver ciertas partes esenciales del bosque y Nagel postula que es necesario

ampliar las bases metodológicas sobre las que se fundó la ciencia natural para reconocer que lo que existe, existe: un nuevo «y sin embargo se mueve» que esperemos no sea anatematizado esta vez por nadie sino, todo lo más, discutido y argumentado con todas las cartas sobre la mesa.

He dicho que Nagel no es antinaturalista y, al advertirlo, no me salgo de lo que él mismo afirma sobre su propia postura. Este libro no defiende ninguna mente suprema en el sentido en que propuso, con escándalo para muchos y tras haber alardeado durante muchos años de ateísmo, el filósofo Anthoy Flew en su obra *Dios existe* (Madrid, Trotta, 2012). Así como en ese libro Flew reconoce haber cambiado su opinión sobre la existencia de un agente intelectual superior aunque no sobre la mortalidad personal, sin embargo, Nagel en *La mente y el cosmos* se declara abiertamente ateo y falto de cualquier tipo de *sensus divinitatis*. No tienen nada que temer los naturalistas no reduccionistas de sus posiciones sobre la divinidad ni que se vaya a unir a los defensores del diseño inteligente o a los argumentos de Plantinga por motivos estrictamente religiosos. Ahora bien, una cosa es que uno se declare ateo y otra distinta que se niegue a entrar en debate y a dar carta de ciudadanía a quien piensa diferente de él, amparándose en el único motivo de que tiene creencias religiosas. Eso sería manifiestamente injusto, como injusto es el trato despectivo que se está dando por parte de la autoproclamada «ciencia seria» a los defensores del diseño inteligente. Lo que Nagel afirma es que la hipótesis científica neo-darwinista que ha derivado filosóficamente en posturas reduccionistas y materialistas (y no las otras no reduccionistas y no materialistas, que también las hay) deja fuera de su ámbito de explicación fenómenos que realmente son contrastables o que para explicarlos deben hacer un conjunto de conjeturas que significan un triple salto mortal en el vacío de cuyo riesgo nadie parece darse cuenta. Y pone el dedo en la llaga, exactamente en la falta de explicación de tres llagas en el ámbito del naturalismo reduccionista: la conciencia, el conocimiento y el valor. Pero vayamos por partes, porque hay en el libro una cuestión previa que los naturalistas reduccionistas suelen criticar a quienes no comparten sus creencias.

Antes de enfrentarnos con esos tres hechos ineludibles e inexplicados por el naturalismo reduccionista, hay otro que afecta a la interpretación

global de la teoría de la evolución y es la espinosa cuestión del origen de la vida. Ese es un tema tan delicado que ha llevado a extremos. Es conocida la postura de Francis Crick de que la vida en la Tierra resulta un milagro y que solo puede ser explicada como efecto de agentes extraterrestres. Pero lo más criticado por los naturalistas es el ataque que algunos teístas hacen a la teoría de la evolución porque es incapaz de explicar el origen de la vida, lo cual haría posible considerar un origen causado por un agente intencional supremo. La respuesta del naturalismo es que la teoría de la evolución parte de la vida y no se debe cuestionar porque deje sin explicar los motivos de su surgimiento. Es cierto que la teoría de la evolución *qua* hipótesis científica parte del hecho de la vida y pretende explicar solo su evolución. Pero, al mismo tiempo, los científicos que la defienden no se preocupan de dejar claro ese mismo hecho y, por omisión, dejan que se extienda el mito materialista de que la vida surge de las estructuras cada vez más complejas de las que se compone la química. Nagel lo que hace es cubrir la responsabilidad no asumida por los teóricos de la evolución diciendo que el origen de la vida desde los constituyentes elementales del universo no es algo de sentido común sino que supone un problema que, de momento, nadie ha solucionado. Además, si no es la propia teoría de la evolución, ¿quién se tiene que encargar de explicar el origen de la vida? ¿No está la propia bioquímica transida de principios evolucionistas? ¿Quién tendría que explicar ese origen sino una teoría *bio*lógica cuyo estándar es, de momento, el neodarwinismo? Por eso puede decirse que, sin dejar de ser naturalista, ya que confía en que alguna vez se dé una solución inmanente a los problemas del orden natural, sin embargo Nagel no es reduccionista porque la cuestión, todavía al menos, es una pregunta sin respuesta. Pero hay algo más, hay un paso hacia adelante más que matiza e incluso altera sustancialmente ese «todavía» que da esperanza a los reduccionistas. Según Nagel las bases metodológicas de la ciencia moderna no podrán explicar nunca esos fenómenos nuevos entre los que se encuentra la vida. Por ello apela a la necesidad de reconfiguración de las bases metodológicas de la ciencia en un campo ampliado y unificado que sea capaz de explicar todos los fenómenos, tanto físicos como biológicos, como también mentales en un

amplio sentido. Lo que pretende Nagel es hacernos conscientes de la necesidad de una nueva revolución científica.

Cambiar esas bases metodológicas no es el objetivo de Nagel. No sé si lo intentará en otro libro; su aspiración ahora es pensar la necesidad de cambiarlas dada la enorme cantidad de fenómenos esenciales que la ciencia natural no puede explicar. Y su pensar libre ha dado como fruto un libro de filosofía que está llamado a convertirse en una referencia necesaria en la bibliografía contemporánea. Da gusto leer un libro que recoge un pensar sincero, en primera persona, si además se hace con rigor. Y eso aunque al final no plantee una elaboración positiva de la nueva ciencia llamémosla, en un sentido distinto del que se la suele llamar desde hace algunos años, postmoderna o, para evitar malentendidos, postmaterialista. Para hacer eso hay que cuestionar el reduccionismo. Nagel lo piensa despacio y lo piensa bien.

Pero no solo está el problema de la vida que, si me apuran, es quizás el más explicable por selección natural. Los problemas principales son las tres «*vulnera*» de la ciencia natural contemporánea: la conciencia, el conocimiento y el valor, que dan título a tres de los capítulos del libro. A este respecto quiero hacer mención al reciente libro del profesor Antonio Diéguez *La evolución del conocimiento* (Madrid, Biblioteca Nueva, 2011). En él defiende un naturalismo coherente que explicaría el surgimiento por evolución de las capacidades superiores del ser humano indicando con toda honestidad qué hay de hecho y qué de conjetura en sus posiciones. Esta obra suscitó, en el año 2012, un intenso debate entre Diéguez y el profesor Francisco José Soler que se encuentra publicado en el número 1 de la revista de estudios interdisciplinares *Naturaleza y libertad* (accesible *on-line* en el portal http://grupo.us.es/naturalezayl/). Diéguez y Soler se han encontrado en alguna que otra ocasión y nunca han defraudado con sus argumentos. La cercanía de sus temáticas a los problemas que plantea Nagel se recoge también en el último libro de Soler *Mitología materialista de la ciencia* (Madrid, Encuentro. 2013) y en la reseña que Diéguez ha publicado del libro de Nagel que estoy presentando en el número 29 de la revista *Ludus vitalis* y que lleva por título «De nuevo, la mente como excepción». La lectura sopesada de esos escritos creo que da aún más valor al libro de

Nagel y aporta elementos de juicio para decidir si es necesario o no abandonar, de una vez por todas, el naturalismo reduccionista.

Pero vayamos ya al análisis de la conciencia. Diéguez considera en su reseña que es el capítulo central, aunque en ese juicio quizás se está dejando llevar por sus propios intereses especulativos. Hay razones para considerar igualmente fundamentales los otros capítulos porque, además de la interioridad y la perspectiva subjetiva de la primera persona, el nacimiento de la búsqueda de la objetividad de la verdad del mundo y de la objetividad de los principios que deben transformar el mundo son novedades tan radicales como la conciencia misma. Es verdad que desde un punto de vista evolutivo el primero sería anterior en el tiempo y más básico temáticamente hablando. Pero, desde el punto de vista de Nagel y de cualquiera que no lleve el evolucionismo más allá de la biología, esos otros aspectos están de alguna forma dejando caer su sombra sobre el mismo tema de la conciencia. Con respecto a la conciencia: ¿qué aporta la perspectiva de la primera persona a la eficacia reproductiva? La respuesta es que, a pesar de lo importante que pueda ser para nosotros nuestra propia intimidad, posiblemente no aporte nada e incluso pueda ser fuente de innumerables sufrimientos. En el ámbito de la conciencia, del darse cuenta subjetivo, se produce el conflicto de las pasiones y, como reacciones naturales o enfermizas, surgen las emociones negativas o autodestructivas. Es, en consecuencia, un ámbito no solo evolutivamente irrelevante, sino claramente desadaptativo. Lo más adaptativo es el comportamiento de una máquina que en cada momento se ajusta a sus circunstancias y dispone de recursos de conducta suficientes como para responder automáticamente a los estímulos del medio. ¿Qué es lo que lleva, entonces, a la evolución a favorecer la aparición de este fenómeno? La respuesta clásica del reduccionista materialista es convertirla en un hecho bruto, es decir, afirmar que no sabemos por qué pero que de hecho ha sido así. Y aquí hay un problema de fondo apuntado por Nagel: acudir al hecho bruto no es dar una explicación de por qué ocurre. Que ha ocurrido ya lo sabemos. Lo difícil es argumentar una explicación en su sentido epistemológico riguroso y el neodarwinismo materialista sencillamente no la tiene. Decir que la conciencia se reduce a elementos físicos y químicos cuya hiper-

complicación con el paso del tiempo produce una emergencia en la forma de un epifenómeno, puede resultar un buen programa de investigación, pero dar esa conjetura el rango de una explicación adecuada resulta excesivo. Lo más urgente de todo, además de dar cuenta de las posibilidades de explicación disponibles, además de entrar a fondo en qué se entiende por explicación y además de un innumerable elenco de cuestiones relevantes, es, simplemente, reconocer que hay un problema. Un problema que el reduccionismo no sabe cómo resolver y que el emergentismo no sabe cómo explicar. Pero ante todo un problema. Después podrá discutirse cuáles son las mejores formas de resolverlo. Como en las crisis de pareja, antes de la terapia, hay que reconocer que hay un conflicto. Y parece que el reduccionismo materialista está todavía en la fase de negación.

Una cuestión similar se plantea con el conocimiento propio de las facultades superiores exclusivas del ser humano: ¿qué aporta el interés que demuestran los seres humanos por la verdad a su supervivencia biológica?, ¿qué más da, para mi supervivencia biológica, si no cruzo una autopista porque sé que lo más probable es que me atropelle un vehículo que si no lo hago porque, según una antigua tradición, da mala suerte y provoca la ira de los dioses? En esa aparente delgada línea se dirime toda la cuestión de lo humano, de lo propiamente humano. Según el neodarwinismo materialista tener un cierto conocimiento válido del mundo es esencial para la supervivencia. Y no le falta razón, siempre y cuando se ciña al conocimiento sensorial y perceptivo que compartimos con otras especies animales. Pero ¿por qué las matemáticas, por qué la música, por qué la filosofía, por qué todo aquello que, en definitiva, nos hace humanos y no tiene una relación directa con la eficacia reproductiva? Desde el punto de vista evolutivo es, simplemente, en término empleado por vez primera por Jay Gould, una «exaptación» o, dicho en forma de expresión, un efecto colateral de la evolución que, por azar, ha devenido en esencial para la supervivencia. Y ciertamente esa es una razón, pero no creo que sea una razón suficiente para el ser humano, como quizás pudiera serlo para el pulgar del panda rojo. Como se dijo antes, para el reduccionismo, lo que no puede explicar o no existe o es algo irrelevante, un efecto secundario. Pero

el problema es que el efecto secundario se ha convertido en lo esencial. Y no se soluciona algo tan chocante creando neologismos *ad hoc*.

Lo anterior se engloba en un punto de vista mucho más amplio, en la creencia de que el universo es inteligible o, siendo justos y parafraseando a Aristóteles, que es indigno del ser humano no alcanzar la verdad a él proporcionada. Pero es que la verdad que está en proporción con el hombre no es la de su supervivencia ni la de su proliferación biológica, sino la de buscar el conocimiento por sí mismo. Esto nos lleva a problemas filosóficos de gran envergadura, en los que Nagel no entra porque caen fuera del propósito de su obra: ¿es la razón connatural con el orden del mundo?, ¿tiene capacidad de acceder a él de forma suficiente? Y la convicción que da como resultado, entre otras cosas, a la ciencia misma es una respuesta positiva a ambas preguntas. Es lógico que a una teoría del conocimiento naturalista neo-darwinista le resulte difícil salir del ámbito del instrumentalismo, del convencionalismo, de la pluralidad de lógicas alternativas o de todos los nuevos términos que se han acuñado para disfrazar el subjetivismo.

Personalmente echo en falta en el libro de Nagel una justificación, aunque fuera somera, de su afirmación de que la razón está en contacto directo con la verdad. En cualquier caso su opción por el realismo científico es clara así como le son inherentes al neodarwinismo caer en distintos tipos de biologismos o psicologismos epistemológicos. Según Nagel la lógica no es producto de la biología, suponer que lo es necesita de una justificación mayor de la que proporciona el neodarwinismo reduccionista. Y eso es así, como un ejemplo más, porque la propia hipótesis científica evolucionista está sometida al juicio que sobre ella realice la propia razón. Nagel tiene todas las de ganar con la lógica realista que está en el fondo de todo ser humano y que le dice que no puede conceder aquiescencia a lo que no le resulta verdadero o, al menos, plausible. Y es que la razón es el último juez en todo esto. Algo así no tiene aspecto de ser en su origen un efecto secundario.

La cuestión del valor de la verdad en la razón teórica guarda estrechas relaciones con el realismo del valor en la acción humana, en la razón práctica. Las conductas derivadas del conocimiento sensorial, perceptivo e

incluso emocional pueden ser explicadas en sus niveles de supervivencia por las leyes de la eficacia reproductiva, ya que hay relación entre lo que permite vivir (comida, bebida, refugio, sexo) y el placer. Sin embargo, ¿a qué obedece el ansia humana de alcanzar un concepto objetivo de bien? ¿Es algo secundario en la vida de los seres humanos? ¿Quiere, en verdad, la gente sobrevivir individual y específicamente agotando el ciclo de nacer-crecer-reproducirse y morir, o eso no les es suficiente y pretenden, más bien, una existencia «auténtica» (en su sentido existencialista) que puede ser muy contraria a lo que se esperaría del fin biológico cumplido de cualquier animal por muy superior que sea? Y eso exige realismo. Nagel lo argumenta directa e indirectamente y llega a la conclusión de que, si el realismo es verdadero, la explicación que da de la moral el reduccionismo materialista que se funda en la hipótesis científica del neodarwinismo es falsa. Como en el caso de las facultades superiores de conocimiento, para el reduccionismo neodarwinista la razón práctica y su objeto más básico, el bien, serían un producto secundario que solo por azar adquiere importancia. El problema aparece cuando uno ve que la cultura en general —de ahí las clásicas críticas de Nietzsche y Freud a la cultura occidental— se han forjado porque la gente no ha rehuido el dolor necesario y ha evitado el placer por valores, *a priori*, muy diferentes de los de la biología. La explicación neodarwinista no considera el bien y el mal objetivo como algo esencial sino como un fenómeno irrelevante para la eficacia biológica. Precisamente ahí se ve que las críticas del irracionalismo a la cultura occidental no solo resultan inadecuadas, sino lisa y llanamente falsas, porque desde sus premisas es imposible iniciar el camino hacia el descubrimiento colectivo del valor. Así se llega a las puertas de otro debate que conectaría con la sociología del conocimiento y en el que también está involucrada la relación entre realismo del valor y subjetivismo, pero que Nagel no toca en su libro.

Ahora bien: si aquello en lo que consistimos esencialmente (vida, conciencia, conocimiento y valor) no es explicado suficientemente por el modelo reduccionista derivado del neodarwinismo, ¿a qué tanto empeño por mantenerlo a toda costa? Una respuesta considerada por Nagel es que no existe otra alternativa mejor. En tal caso, lo que habría que hacer es

buscarla y no aferrarse a un clavo ardiendo. Otra explicación posible haría más lógico el temor del naturalista: abandonar el neodarwinismo significaría abandonar el naturalismo. Pero no parece que eso sea así. Cabe seguir investigando en líneas de explicación evolucionistas no reduccionistas. Quién sabe si en el futuro no tendremos mejores argumentos naturalistas que los que tenemos ahora. Bien es cierto que en la búsqueda pueden aparecer alternativas no naturalistas. Y puede que ese, más que el anterior, sea el auténtico motivo de aguantar con la espada de Damocles encima.

¿Qué alternativas puede haber al materialismo reduccionista? A los dos problemas fundamentales generales de toda ciencia señalados en el libro, el constitutivo y el histórico, cabrían tres posibles respuestas según Nagel: una causal (eficiente), otra teleológica y otra intencional. A la que más atención presta es a la primera, puesto que está más en sintonía con el paradigma de la revolución científica moderna. A la que concede más valor es a la teleológica, lo que supondría un retorno a la idea aristotélica de causas no dirigidas en una única dirección sino hacia una pluralidad de formas cuya virtualidad consiste en que el fin alumbra el principio mismo. La explicación intencional prácticamente ni la considera, porque haría intervenir una mente divina en el proceso que, o bien creó la materia y las leyes y después dejó el universo en paz, explicando así el problema constitutivo, o interviene varias veces en la historia del proceso, dando razón sus intervenciones del problema histórico. Para entender bien el panorama actual de estas tres respuestas, remito al libro de Juan Arana *Los sótanos del Universo* (Madrid, Biblioteca Nueva, 2012) y al dossier sobre esa obra que acaba de publicarse en Enero de 2014 como número 3 de la revista *Naturaleza y libertad*. Creo que esas lecturas complementarias, junto con la del libro de Nagel, dispensan de más precisiones al respecto.

A la luz de lo expresado por Nagel se comprende que se pueda ser teísta y discrepar del diseño inteligente. También se puede ser creyente y optar como científico por un naturalismo no reduccionista. Las disputas filosóficas no tienen por qué distraer a los científicos experimentales de sus experimentos. Pero la ciencia también interesa al hombre común y, lo que es más, es conveniente que la conozca y tome postura crítica ante ella. Lo

único que prohíbe la coherencia es ser creyente y materialista al mismo tiempo. Eso lo saben bien los materialistas y por ello se arman hasta los dientes de argumentos contra la religión. La postura del creyente y del científico antirreduccionista es dar argumentos válidos contra el materialismo. Lo peor que se puede hacer por la verdad en la que uno cree es defenderla con malos argumentos o, peor incluso, no conocer por desidia los argumentos existentes a favor de las propias creencias. El libro de Nagel no pone bases para sostener el teísmo, como hace el libro de Flew citado con anterioridad, pero proporciona un número considerable de dudas que afectan al materialismo y suponen un inicio prometedor para gestar un nuevo paradigma científico que explique los hechos desde el realismo.

La pregunta sería en definitiva: ¿hay que tomarse en serio la propuesta de Nagel de una nueva revolución científica? Si el objetivo que pretende la ciencia es explicar el conjunto de los fenómenos, entonces no queda otra posibilidad, especialmente si dentro del conjunto de la realidad se quiere incluir los fenómenos de la conciencia, el conocimiento objetivo y el valor, cuyos orígenes no pueden ser explicados de forma suficientemente completa por el neodarwinismo reduccionista. Ciertamente las posibles soluciones que Nagel propone para un fundamento ampliado y unificado de la ciencia son altamente especulativas y caen dentro del campo de la filosofía más que de la ciencia. Pero todo el mundo sabe que en el conocimiento de frontera, en los programas que mueven a la investigación, lo empírico y lo especulativo se entremezclan de tal forma que es muy difícil separar lo estrictamente establecido de la interpretación filosófica. Ciencia y filosofía están unidas en los planteamientos de alcance y los resultados se obtienen tanto mediante la investigación empírica como por medio de la crítica racional. Si la revolución científica amputó la mente del orden de sus prioridades tiene que integrarla nuevamente para poder explicarla de verdad. Es dudoso que postular entidades proto-psíquicas tan básicas como las partículas elementales sea una buena solución. Es probable que la idea nos lleve más bien a un tipo de entidades proto-físico-psíquicas cuya diversificación explicaría la posibilidad de que en la historia de la evolución aparezca lo mental. Se trata, por lo tanto, de indagar las condiciones de posibilidad de esa aparición. Este hecho viene a poner en su

justo lugar una de las soluciones que da el profesor Diéguez en su reseña del libro de Nagel: los fenómenos nuevos serían explicados por lo que él llama un «*feedback* fructífero». Según este planteamiento, el conocimiento sensible se enfrentaría a los resultados de su propia acción y los evaluaría para ajustar sus respuestas en una nueva ocasión; además se valorarían también las respuestas de los otros tomándose como propias y así, reevaluando, resultaría el conocimiento objetivo. Creo que esa interpretación sigue sin explicar la novedad de la acción por motivos ajenos a la eficacia biológica. ¿Las leyes de la evolución biológica superan las leyes de la evolución biológica? Eso es lo que se afirma, pero creo que no se explica suficientemente: la novedad debe ser posible y entonces hay que considerar las leyes evolutivas como un paso intermedio entre el punto de partida y el de llegada, pero no como los únicos mecanismos de explicación posible precisamente porque hay que explicar el punto de partida y comprender el punto de llegada que, de otra forma, quedarían sumidos en el misterio. Explicar el surgimiento de las tres *vulnera* tan solo por el azar más la selección natural tiene, al menos, dos consecuencias que resultan indeseables: no se puede dar cuenta de ellas más que como hechos brutos, es decir, han ocurrido porque han ocurrido y ya está, lo cual implica que en el fondo quedan inexplicadas si seguimos un concepto riguroso de explicación tal como defiende Nagel en su obra. En segundo lugar, el tipo de realismo al que conduce ese planteamiento no logra salir del subjetivismo o de un convencionalismo cuyas últimas razones no superan los postulados del utilitarismo. El subjetivismo utilitarista es lo máximo a lo que puede llegar una explicación neodarwinista de la vida práctica individual y social. Cuestionaría el realismo del valor, un realismo que no puede surgir como resultado del azar sino que, de alguna manera, está contenido de antemano en la propia estructura del universo como totalidad inteligible.

La obra de Nagel no rechaza que el neodarwinismo tenga un apoyo empírico suficiente para explicar los hechos biológicos; lo que cuestiona es que su variante materialista pretenda ser una explicación totalizadora de los fenómenos del universo. En ese sentido puede decirse que la versión materialista del neodarwinismo es falsa. Para que no lo fuera debería

reconocer sus propios límites explicativos o, al menos, reconocer que, como teoría de la totalidad, es tan solo uno más de los programas de investigación que pueden surgir del neodarwinismo. El análisis crítico del reduccionismo es, sin duda, lo más logrado del trabajo de Nagel. Cumple sobradamente los objetivos de señalar los defectos insalvables de una interpretación de la teoría de Darwin que ha hipertrofiado sus propósitos originales convirtiéndose en teoría del todo y, que, para que recupere la salud y deje espacio libre a planteamientos alternativos rigurosos, es necesario someterla a un poderoso ejercicio de crítica. Nagel abre la puerta a la búsqueda de alternativas al modelo reduccionista. Tendrán que ser otros quienes se ocupen de resolver los problemas que aparezcan al explorar esas otras alternativas. Pero lo que está fuera de duda es que, en el estado actual de la ciencia y de la divulgación científica, se impone que la filosofía realice la *pars destruens* que Nagel ha llevado a cabo. Ahora se presenta un futuro más despejado para trabajar con más libertad. La ciencia nunca pretende impedir que se siga pensando lo cual, sin embargo, es más propio de la ideología. La ideología está ocupando el lugar de la ciencia e impide pensar un conjunto de alternativas que a todas luces se revelan como posibles. El futuro queda abierto con la obra de Nagel. Sumérjase el lector en sus argumentos y sopéselos. Encontrará frescura, rigor argumentativo y liberación de prejuicios.

Es de justicia que acabe esta presentación agradeciendo a la editorial Biblioteca Nueva y a la Asociación de Filosofía y Ciencia Contemporánea el esfuerzo que han realizado para que este libro viera la luz en castellano. Tengo que agradecer también a Juan Arana y Concepción Diosdado las revisiones que han hecho tanto del texto de la obra de Nagel como de la presentación. Sus comentarios y sugerencias han sido muy valiosos para corregir imperfecciones y errores. En la medida en que soy consciente los he asumido todos. De todas formas los exonero de la culpa de cualquier malinterpretación y fallo que haya podido cometer y de los que yo únicamente quiero asumir su responsabilidad.

Francisco Rodríguez Valls

A Anne

PREFACIO

Las dos personas que, de formas muy distintas, más han influido en los pensamientos expuestos en este libro son Sharon Street y Roger White. También fueron muy instructivas y estimulantes las discusiones del grupo de investigación sobre ciencia y religión organizado por el Instituto de Filosofía de Nueva York con el apoyo de un *Distinguised Achievement Award* de la Fundación Mellon. Agradezco a la Fundación Mellon haberlo hecho posible. El grupo, que se reunió desde 2006 hasta 2009, congregó a estudiantes de grado y de posgrado del departamento de filosofía de la Universidad de Nueva York con participantes regulares y ocasionales de otras universidades y otros campos. Street y White fueron miembros de ese grupo. También me gustaría mostrar mi agradecimiento en particular a Paul Boghossian, Laura Franklin-Hall, Philip Kitcher, Matthew Kotzen, H. Allen Orr, Alvin Plantinga, Elliot Sober y a Michael Strevens. Sober leyó también el manuscrito del libro para Oxford University Press e hizo sugerencias útiles. He presentado algunos de los materiales en el Coloquio sobre Filosofía Legal, Política y Social que Ronald Dworkin y yo hemos dirigido durante muchos años y le estoy agradecido a él y a los otros participantes por su ayuda. En vistas de lo heterodoxo del resultado espero que mi gratitud no les ofenda.

Durante la redacción del libro recibí una ayuda a la investigación de la *Filomen D’Agostino y Max E. Greenberg Faculty Research Fund* de la Escuela de Leyes de la Universidad de Nueva York.

Thomas Nagel
Nueva York, octubre de 2011

INTRODUCCIÓN

El objetivo de este libro es argumentar que el problema mente-cuerpo no es un problema localizado en un ámbito de conocimiento particular que solo tiene que ver con la relación entre mente, cerebro y conducta en los organismos animales vivos, sino que realmente penetra nuestra comprensión del cosmos entero y de su historia. Las ciencias físicas y la biología evolutiva no pueden ser ajenas a él y creo que reconocer la dificultad del problema debe, en última instancia, cambiar nuestra concepción del lugar de las ciencias físicas en la descripción del orden natural.

Una de las tareas legítimas de la filosofía es investigar los límites de las formas más desarrolladas y más exitosas del conocimiento científico contemporáneo. Puede resultar frustrante reconocerlo, pero estamos simplemente en el punto de la historia del pensamiento humano en el que estamos y nuestros sucesores harán descubrimientos y desarrollarán formas de conocimiento que ni siquiera hemos soñado. Los humanos tenemos la esperanza de llegar a un punto final, pero la humildad intelectual requiere que resistamos la tentación de creer que los instrumentos que tenemos ahora son suficientes para comprender el universo como un todo. Sea el que sea quien se ocupe de ella, señalar sus límites es una tarea filosófica más que una parte de la búsqueda interna de la ciencia, aunque podamos esperar que, si se reconocen los límites, ello pueda finalmente conducirnos al descubrimiento de nuevas formas de comprensión científica. Los científicos son muy conscientes de cuánto desconocen, pero este problema es de una clase diferente: reconocer no solo los límites de lo que de hecho comprendemos, sino intentar reconocer aquello que puede o no puede ser comprendido por medio de ciertos métodos existentes.

Mi meta es obtener una imagen del mundo comprehensiva y especulativa que se obtenga por la extrapolación de algunos de los descubrimientos de la biología, la química y la física; alcanzar una particular *Weltanschauung* naturalista que postula una relación jerárquica entre los objetos de estas ciencias y que busca, en principio, una explicación completa de todo el universo capaz de lograr su unificación. Tal concepción del mundo no es una condición necesaria para la práctica de ninguna de las ciencias y, aceptarla o no, no tendría efecto sobre la mayor parte de la investigación científica. Por lo que sé, la mayoría de los científicos prácticos no tienen por qué tener opinión alguna sobre los problemas cosmológicos dominantes a los cuales el reduccionismo materialista ofrece una respuesta. Su investigación concreta y hallazgos sustantivos no dependen en general, ni tampoco implican, esta o aquella respuesta a tales cuestiones. Pero entre los científicos y filósofos que expresan puntos de vista sobre el orden natural como un todo se asume ampliamente que el materialismo reductivo es la única posibilidad seria a considerar[1].

El punto de partida de mi planteamiento es el fracaso del reduccionismo psicofísico, una postura en filosofía de la mente que se mueve especialmente por la esperanza de mostrar cómo las ciencias físicas podrían en principio proporcionar una teoría de la totalidad. Si esa esperanza fuera irrealizable surgiría la pregunta de si otra comprensión más o menos unificada podría hacerlo con el cosmos entero tal como lo conocemos. Entre los candidatos tradicionales para entender comprehensivamente la relación de la mente con el mundo físico creo que el peso de la evidencia favorece a alguna forma de monismo neutral sobre las alternativas tradicionales de materialismo, idealismo o dualismo. Lo que me gustaría hacer es explorar las posibilidades que son compatibles con lo que sabemos, en particular lo que sabemos sobre cómo la mente y todo lo relacionado con ella dependen de la aparición y desarrollo de los organismos vivos como resultado de la física del universo, la química y, después, la evolución biológica. Argumentaré que esos procesos deben ser repensados a la luz de lo que ellos mismos han producido si es que el reduccionismo psicofísico es falso.

El argumento del fracaso del reduccionismo psicofísico es filosófico, pero creo que hay razones empíricas independientes para dudar de la verdad del reduccionismo en biología. En biología el reduccionismo físico-químico es el punto de vista ortodoxo y cualquier oposición a él se considera no solo científica sino políticamente incorrecta. Pero durante mucho tiempo he encontrado difícil de creer la explicación materialista de cómo nosotros y el resto de los organismos llegamos a existir, incluyendo la versión estándar de cómo funciona el proceso evolutivo. Cuantos más detalles aprendemos sobre las bases químicas de la vida y lo intrincado del código genético, más falta de crédito resulta tener la explicación histórica estándar[2]. Esta es solo la opinión de un lego que lee toda la bibliografía que explica la ciencia contemporánea a los no especialistas. Quizás esa bibliografía presenta la situación con una simplicidad y confianza que no refleja el pensamiento científico más sofisticado en esas áreas. Pero me parece que, tal como se suele presentar, la corriente ortodoxa sobre el orden cósmico es el producto de conjeturas gobernantes que no tienen fundamento y que van en contra del sentido común.

Quiero defender en este libro la espontánea reacción de incredulidad de un no especialista al reduccionismo de la explicación neodarwinista del origen y evolución de la vida[3]. Es, *prima facie*, altamente implausible que la vida como la conocemos sea el resultado de una secuencia de accidentes físicos conjugados con el mecanismo de la selección natural. Espero abandonar la simpleza de esta respuesta inicial, no por una explicación físico/química completamente terminada, sino por una alternativa que realmente sea un esquema de explicación que esté apoyado por algunos ejemplos. Lo que falta, a mi entender, es un argumento creíble de que el asunto tiene una probabilidad, y no pequeña, de ser verdad. Hay dos preguntas. Primero, dado lo que se sabe sobre las bases químicas de la biología y la genética, ¿cuál es la probabilidad de que formas vivas que se auto-replican puedan haber llegado a la existencia espontáneamente al principio de la Tierra solamente a través de la actuación de las leyes de la física y la química? La segunda pregunta gira en torno a las fuentes de variación del proceso evolutivo que se puso en marcha una vez que empezó la vida: en el tiempo geológico disponible desde que aparecieron las

primeras formas de vida sobre la tierra, ¿cuál es la probabilidad de que, como consecuencia de un accidente físico, haya ocurrido una secuencia de mutaciones genéticas viables que fuera suficiente para permitir que la selección natural produjera los organismos que existen de hecho?

Hay mucha menos certeza en la comunidad científica sobre la primera pregunta que sobre la segunda. Mucha gente piensa que será muy difícil continuar defendiendo una explicación reduccionista del origen de la vida, pero la mayoría no tiene duda de que la variación genética accidental es suficiente para avalar la evolución por selección natural una vez que los organismos que se reproducen han aparecido. Sin embargo, en tanto que los problemas se refieren a eventos muy específicos ocurridos en largos períodos históricos en un pasado distante, la evidencia disponible es muy indirecta y las conjeturas deben desempeñar una parte importante en la explicación de esos problemas. Mis dudas no se basan en creencias religiosas o en una creencia en una alternativa plenamente definida. Es solo una creencia en que la evidencia científica disponible, a pesar del consenso de la opinión científica, no nos exige racionalmente dar prioridad a esa explicación sobre la reticencia que le causa al sentido común. Esto es especialmente verdad en lo que se refiere al origen de la vida.

El mundo es un lugar asombroso y la idea de que tenemos en nuestra posesión los instrumentos básicos necesarios para comprenderlo no es más creíble hoy de lo que fue en los días de Aristóteles. Que nos ha producido a ti, a mí y al resto de nosotros es lo más asombroso de todo. Si la investigación contemporánea en biología molecular deja abierta la posibilidad de dudas legítimas sobre una explicación completamente mecanicista del origen y evolución de la vida, dependiente solamente de las leyes de la química y la física, ello puede combinarse con el fracaso del reduccionismo psicofísico para sugerir que hay principios de un tipo diferente que están operando en la historia de la naturaleza, principios de crecimiento del orden que son, en su forma lógica, teleológicos más que mecanicistas. Me doy cuenta de que tales dudas sonarán a muchos como una extravagancia, pero eso es porque casi todos en nuestra cultura laica han sido compelidos a considerar como sacrosanto el programa de

investigación reduccionista con el argumento de que cualquier otro no sería ciencia.

Mi proyecto tiene la forma habitual de intentar encontrar un conjunto de condiciones que parece imposible que se den unidas. Además del antirreduccionismo, dos restricciones más son importantes: primero, poseer una creencia de que ciertas cosas son tan relevantes que tienen que ser explicadas como no accidentales si queremos alcanzar un conocimiento real del mundo; segundo, tener el ideal de descubrir un orden natural simple que unifique todo sobre la base de un conjunto de elementos y principios comunes, un ideal al que las formas inevitablemente muy incompletas de nuestro conocimiento deberían aspirar a pesar de todo. El dualismo cartesiano rechaza esta segunda aspiración y los programas reduccionistas tanto materialistas como idealistas han fracasado en el intento de llevarlo a cabo. Esa concepción unificada es también incompatible con una clase de teísmo que explica ciertas formas del mundo natural por intervención divina, la cual no es parte del orden natural.

Los grandes avances en las ciencias físicas y biológicas fueron posibles por la exclusión de la mente del mundo físico. Ello permitió una comprensión cuantitativa del mundo según leyes físicas formuladas matemáticamente y, en cuanto tales, expresadas intemporalmente. Pero en algún momento será necesario recomenzar sobre un modelo más comprehensivo que incluya la mente. Un modelo tal tendrá tanto una dimensión histórica como una intemporal. La idea de que la comprensión histórica es parte de la ciencia se ha hecho familiar mediante la transformación de la biología por la teoría evolucionista. Pero más recientemente, con la aceptación del *big bang*, la cosmología también se ha convertido en una ciencia histórica. La mente, como un desarrollo de la vida, debe ser incluida en esa explicación como el escenario más reciente de esta larga historia cosmológica y su aparición deja caer su sombra sobre el proceso entero y sobre los constitutivos y principios de los que depende el proceso.

El problema es si podemos integrar esta perspectiva con la de las ciencias físicas, ya que estas han sido desarrolladas para explicar un universo amental. La comprensión de la mente no puede quedar en una

explicación subjetiva, puesto que la mente es producto, en parte, de un proceso físico; pero, por la misma razón, el aislamiento de la ciencia física y su exigencia de completitud tiene que terminar a la larga. Y eso plantea la pregunta: ¿en qué medida la forma reductiva que es tan central a la ciencia física contemporánea sobrevivirá a la transformación? Si la física y la química no pueden explicar completamente la vida y la conciencia, ¿cómo se combinará su inmenso cuerpo de doctrina con otros elementos en una concepción ampliada del orden natural que pueda albergar todas estas cosas?

Como he dicho, plantear dudas sobre la explicación reduccionista de la vida va en contra del consenso científico dominante, pero ese consenso se enfrenta a problemas que creo que no son tomados con bastante seriedad, tanto en lo que se refiere a la evolución de las formas de vida a través de la mutación accidental y la selección natural como en lo que se refiere a la formación desde la materia inerte de sistemas físicos capaces de tal evolución. Cuanto más aprendemos lo intrincado que es el código genético y su control de los procesos químicos de la vida tanto más difíciles resultan estos problemas.

De nuevo: respecto a la evolución, el proceso de la selección natural no puede dar razón de la historia real sin un suministro adecuado de mutaciones viables y creo que es un problema abierto si eso puede haber ocurrido en el tiempo geológico como resultado de un simple accidente químico, sin la intervención de algunos otros factores que determinen y restrinjan la formas de variación genética. Ya no es legítimo, simplemente, imaginar una secuencia de fenotipos desenvolviéndose gradualmente como si su aparición a través de mutaciones en el ADN no fuera problemática, como Richard Dawkins hace con la evolución del ojo[4]. En lo referente al origen de la vida, el problema es mucho más difícil en tanto que no disponemos para explicarlo de la opción de la selección natural. Y la aparición del código genético —una arbitraria trasposición de secuencias de nucleótidos en aminoácidos junto con los mecanismos que pueden leer el código y cumplir sus instrucciones— parece que resiste particularmente el revelarse como probable si se da solo la ley física[5].

Al pensar sobre estos problemas me han motivado las críticas de la imagen del mundo científico dominante que se han hecho desde una dirección diferente: el ataque al darwinismo realizado en años recientes desde una perspectiva religiosa por los defensores del diseño inteligente. Aunque escritores como Michael Behe y Stephen Meyer están movidos, al menos en parte, por sus creencias religiosas, los argumentos empíricos que ofrecen en contra de la probabilidad de que el origen de la vida y su historia evolutiva puedan ser completamente explicados por la física y la química son de gran interés por ellos mismos[6]. Otro escéptico, David Berlinski, ha destacado con viveza estos problemas sin referencia a la inferencia del diseño[7]. Incluso si uno no concluye en la alternativa de una explicación que requiera la acción de un diseñador, los problemas que estos iconoclastas plantean al consenso científico ortodoxo deberían tomarse seriamente[8]. No se merecen el desprecio con el que se encuentran comúnmente. Es manifiestamente injusto.

Los que han criticado con seriedad estos argumentos han mostrado que hay formas de oponerse a la conclusión del diseño; pero la fuerza de la parte negativa de la posición del diseño inteligente, a saber, sus dudas sobre la probabilidad del punto de vista reductivo ortodoxo dada la evidencia disponible, no me parece que haya sido destruida con esas críticas[9]. Al menos el problema se debería considerar abierto. Para todo aquel que esté interesado en los fundamentos de este juicio solo puedo recomendarle una lectura cuidadosa de algunos de los principales abogados de ambos lados del contencioso, con especial atención a lo que ha sido establecido por los críticos del diseño inteligente. Independientemente de lo que uno pueda pensar sobre la posibilidad de un diseñador, la doctrina prevalente, de que la aparición de la vida de la materia inerte y su evolución a través de la mutación accidental y la selección natural hasta sus formas presentes no presupone nada sino la operación de la ley física, no puede considerarse como irrefutable. Es una suposición que gobierna el proyecto científico más que una hipótesis científica bien confirmada.

Confieso que mi suposición de no considerar posible la alternativa del diseño como una opción real carece de fundamento. Me falta el *sensus divinitatis* que capacita —en verdad compele— a mucha gente a ver en el

mundo la expresión de un propósito divino tan naturalmente como ven en una cara sonriente la expresión del sentimiento humano[10]. Por lo tanto mis especulaciones sobre una alternativa a la física como teoría del todo no invocan en absoluto a un ser trascendente sino que se orientan a resolver las complicaciones surgidas dentro de la inmanencia del orden natural. Esa también sería una explicación más unificadora que la hipótesis del diseño. Discrepo de los defensores del diseño inteligente en su suposición, que comparten con sus oponentes, de que la única alternativa naturalista es una teoría reduccionista fundada en las leyes físicas del tipo a las que estamos acostumbrados. A pesar de eso, creo que los defensores del diseño inteligente merecen nuestra gratitud por plantar cara a una concepción científica del mundo que despierta pasión en sus seguidores en parte por el hecho de que piensan que nos libera de la religión.

Esa comprensión del mundo está madura para cambiar a pesar de los grandes logros del materialismo reductivo, que presumiblemente continuará por largo tiempo siendo nuestra principal fuente de comprensión concreta y de control del mundo que nos rodea. Plantear, como quiero, que hay muchas cosas que no puede explicar no es ofrecer una alternativa. Pero reconocer esos límites es una condición previa para buscar alternativas o, al menos, para estar abiertos a su posibilidad. Y eso puede querer decir que algunas direcciones de búsqueda de la forma materialista de explicación llegarán a verse como vías muertas. Si la aparición de los organismos conscientes en el mundo se debe a principios de desarrollo que no se derivan de las leyes atemporales de la física, esa puede ser también una razón para ser pesimistas sobre las explicaciones puramente químicas del origen de la vida.

CAPÍTULO 1

EL ANTIRREDUCCIONISMO Y EL ORDEN NATURAL

1

El conflicto entre el naturalismo científico y las diversas formas de antirreduccionismo es un problema esencial de la filosofía reciente. De un lado tienen la esperanza de que todo pueda ser explicado al nivel más básico por las ciencias físicas, ampliadas para incluir a la biología[11]. Del otro lado, hay dudas sobre si la realidad de ámbitos de nuestro mundo tales como la conciencia, la intencionalidad, el significado, el propósito, el pensamiento y el valor pueden encontrar sitio en un universo que consiste en ese nivel más básico solo en hechos físicos, hechos, aunque sean sofisticados, del tipo mostrado por las ciencias físicas.

Usaré los términos «materialismo» o «naturalismo materialista» para referirme a uno de los lados de este conflicto y «antirreduccionismo» para referirme al otro. Reconozco que los términos son más bien aproximados. Los intentos de defender la imagen del mundo materialista como una explicación potencialmente completa de lo que hay adquiere muchas formas, y no todas ellas envuelven reducción en su sentido ordinario, tal como ocurre con el análisis de los conceptos mentales en términos conductistas o la identificación científica de los estados mentales con estados cerebrales. Muchos naturalistas materialistas no describirían sus enfoques como reduccionistas. Pero para los que dudan de la suficiencia de ese punto de vista sobre el mundo, los diferentes intentos de incorporar dentro de él la mente y otros fenómenos relacionados aparecen todos como intentos de reducir la verdad completa de la realidad a principios que no son

suficientemente ricos para ese propósito. De ahí que esa postura de resistencia pueda concebirse como antirreduccionismo.

Las dudas antirreduccionistas se plantean generalmente de forma negativa. La conclusión a la que nos invitan es a que hay algunas cosas que las solas ciencias físicas no pueden explicar completamente. Se pueden necesitar otras formas de comprensión o, quizás, hay más realidad de la que puede describir la física incluso si se la desarrolla completamente. Si la reducción fracasa, ello revela un límite del alcance de las ciencias físicas que debe ser completado por algo que explique los elementos que se echan en falta. Pero el problema que señalan requiere mucho más. Si uno duda de la reductibilidad de lo mental a lo físico, e igualmente de todo lo demás que acompaña a lo mental, como el valor y el significado, entonces hay razones para dudar de que pueda bastar un materialismo reductivo incluso en biología y, por tanto, razones para dudar de que el materialismo pueda dar cuenta adecuada del mundo físico. Quiero explorar la posibilidad de tales dudas y considerar si se puede vislumbrar algo positivo tras ello.

Nosotros y otras criaturas con vidas mentales somos organismos y parece que nuestras capacidades mentales dependen de nuestra constitución física. De ahí se deduciría que lo que explica la existencia de organismos como nosotros debería también explicar la existencia de la mente. Pero si lo mental no es meramente físico no puede ser completamente explicado por la ciencia física. Y entonces, como defenderé, es difícil evitar la conclusión de que los aspectos de nuestra constitución que comportan lo mental no pueden ser explicados completamente tampoco por la ciencia física. Si la biología evolutiva es una teoría física —como se la considera generalmente — entonces no puede dar cuenta de la aparición de la conciencia ni de otros fenómenos que no son físicamente reductibles. Así, si la mente es un producto de la evolución biológica —si los organismos con vida mental no son anomalías milagrosas sino parte integral de la naturaleza— entonces la biología no puede ser exclusivamente ciencia física. Esa sola posibilidad nos abre a una concepción del orden natural muy diferente del materialismo ya que convierte a la mente en algo central más que en un efecto colateral de la ley física.

Es obvio que la conclusión de los argumentos antirreduccionistas contra el materialismo no puede permanecer para siempre en una forma negativa. Incluso si el dominio del naturalismo materialista estuviera cercano a su fin, necesitaremos alguna idea de lo que pueda reemplazarlo. Uno de los argumentos que dirigen los diferentes programas reduccionistas sobre la mente, el valor y el significado, a pesar de su inherente falta de plausibilidad, es la carencia de una alternativa comprehensiva. Parecería que la única forma de aceptar el argumento contra la reducción sería añadir ingredientes extra como los *qualia*, los significados, las intenciones, los valores, las razones, las creencias y los deseos al orden matemático magnificentemente unificado del universo físico. Pero esto no satisface el deseo de una comprensión general de cómo las cosas deben unificarse. Una alternativa genuina al programa reduccionista requerirá una explicación de cómo la mente y todo lo que la acompaña es inherente al universo.

Estoy únicamente dando la vuelta a un argumento común para cuestionar sus premisas. El materialismo requiere reduccionismo; por tanto, el fracaso del reduccionismo requiere una alternativa al materialismo. Mi propósito no es tanto argumentar contra el reduccionismo como investigar las consecuencias de rechazarlo, presentar el problema más bien que proponer una solución. El naturalismo materialista lleva al reduccionismo porque niega la realidad de todas esas cosas tan comunes que no son físicas a primera vista. Pero si no se dispone de una reducción plausible y si negar la realidad de lo mental continúa siendo inaceptable, eso sugiere que la premisa original, el naturalismo materialista, es falso, y no solo por razones nimias. Quizás el orden natural no es exclusivamente físico. O, quizás, en el peor de los casos, no hay orden natural comprehensivo en el que todo esté unido, solamente formas inconexas de comprensión. Pero, independientemente del resultado al que se llegue, debemos iniciar una búsqueda desde una concepción más amplia de lo que hay que comprender para intentar dar un sentido completo al mundo natural.

La convicción que me guía es que la mente no es solo una eventualidad, un accidente o un añadido, sino un aspecto básico de la naturaleza. Aparte de los argumentos antirreduccionistas en filosofía de la mente, hay apoyo independiente para dar pasos hacia una concepción ampliada de la realidad en uno de los supuestos esenciales de la ciencia: la ciencia se mueve por la suposición de que el mundo es inteligible. Esto es, el mundo en el que nos encontramos y sobre el que la experiencia nos da información, puede ser no solamente descrito sino comprendido. Esa creencia está detrás de cada búsqueda de conocimiento, incluso de las búsquedas que terminan en fiasco. En las ciencias naturales, tal y como se han desarrollado desde el siglo XVII, existe la creencia de la inteligibilidad, que ha conducido a descubrimientos extraordinarios confirmados por predicciones y experimentos, de un orden natural oculto que no puede ser observado solo por la percepción humana. Sin la creencia de un orden inteligible subyacente, que antecede con mucho a la revolución científica, estos descubrimientos no podrían haberse logrado.

¿Qué explica este orden? Una respuesta sería que nada lo hace: la explicación culmina con el orden mismo que la creencia de inteligibilidad nos ha capacitado simplemente para descubrir. Quizás un nivel de orden pueda explicarse por un nivel todavía más profundo, como ha pasado repetidamente en la historia de la ciencia. Pero, desde este punto de vista, la comprensión del mundo alcanzará finalmente un punto en el que no haya más que decir excepto: «Esto es exactamente como son las cosas».

No estoy dispuesto a pensar que una ciencia así vaya a tener éxito. Me parece que uno no puede comprender realmente el punto de vista científico a menos que uno asuma que la inteligibilidad del mundo, tal y como es descrita por las leyes que la ciencia ha descubierto, es ella misma parte de una explicación más profunda de por qué las cosas son como son. Así, cuando preferimos una explicación a otra contando con los mismos datos porque es más simple y hace menos suposiciones arbitrarias, no es solo por una preferencia estética: es porque pensamos que la explicación que da mayor comprensión es más probable que sea verdad, solo por esa razón.

Esta creencia es una forma del principio de razón suficiente: todo en el mundo puede de alguna manera ser entendido y si muchas cosas, incluso la

más universal, parecen inicialmente arbitrarias es porque hay más adelante cosas que no conocemos que explicarán por qué, después de todo, no son arbitrarias.

El punto de vista de que la inteligibilidad racional está en la raíz del orden natural me convierte, en un sentido amplio, en un idealista. No en un idealista subjetivo en tanto que no equivale a declarar que toda la realidad es en última instancia apariencia, sino un idealista objetivo en la tradición de Platón y también quizás de ciertos postkantianos como Schelling y Hegel, a los que generalmente se les llama idealistas absolutos. Sospecho que hay un deje de esta clase de idealismo en todo científico teórico: el empirismo puro no basta.

La inteligibilidad del mundo no es ningún accidente. La mente, desde este punto de vista, está doblemente referida al orden natural. La naturaleza es tal como da lugar a los seres conscientes con mentes y es tal como es comprensible a tales seres. Por último y por tanto, tales seres deberían comprenderse a ellos mismos. Y estos son rasgos fundamentales del universo, no productos secundarios de desarrollos contingentes cuya verdadera explicación se daría en términos que no hacen referencia a la mente.

3

La pregunta más amplia que imbuye toda la ciencia natural es también la pregunta más amplia de la filosofía: ¿en qué forma o formas es el mundo inteligible? Claramente la ciencia natural es una de las formas más importantes de desvelar la inteligibilidad. Pero a pesar de los grandes logros de las ciencias naturales en su forma presente, es importante tanto para la ciencia misma como para la filosofía preguntar cuánto de lo que existe pueden hacer inteligible las ciencias físicas, cuánta inteligibilidad del mundo puede ser subsumida bajo leyes universales formulables matemáticamente que gobiernen el orden espacio-temporal. Si hay límites a lo que alcanza la ciencia entendida de esta manera, ¿hay otras formas de comprensión que puedan volver inteligible lo que la ciencia física no explica?

Pero primero deberíamos considerar el punto de vista de que no hay tales límites, que la ley física tiene los recursos para explicar todo, incluyendo la doble relación de la mente con el orden natural. La inteligibilidad (para nosotros) que hace posible la ciencia es una de las cosas que necesitan explicación. La estrategia es intentar ampliar la imagen materialista del mundo de manera que incluya tal explicación, hacer por tanto posible que la inteligibilidad física del mundo se cierre sobre sí misma. Según este tipo de naturalismo, la existencia de mentes para las que el mundo es científicamente inteligible es ella misma científicamente explicable como un efecto colateral biológico altamente específico de orden físico.

La historia va de esto: no hay necesidad de una forma ampliada de comprensión, en su lugar la historia del conocimiento humano nos da razón para creer que hay un único camino en el que el orden natural es inteligible, a saber, a través de la ley física: todo lo que existe y todo lo que ocurre puede en principio ser explicado por las leyes que gobiernan el universo físico. Se admite que no podemos captar el orden natural en toda su manifestación porque es demasiado complejo y que, por tanto, necesitamos formas más específicas de comprensión para fines prácticos. Pero podemos intentar descubrir los principios universales que gobiernan los elementos de los cuales todo se compone y de los que toda complejidad espacio-temporal observable es una manifestación. Estos principios son las leyes matemáticamente establecidas de la física elemental, la cual describe las fuerzas fundamentales y las partículas u otras entidades y sus interacciones, al menos hasta que se descubra un nivel todavía más básico. La descripción más sistemática posible de un universo material extendido en el espacio y en el tiempo es, por tanto, la ruta a la explicación más básica de todo.

La física y la química han buscado este objetivo con un éxito espectacular. Pero el gran paso adelante para el progreso de la concepción materialista hacia el ideal de completitud fue la teoría de la evolución, más tarde reforzada y enriquecida con la biología molecular y el descubrimiento del ADN. La teoría evolutiva moderna ofrece una imagen general de cómo la existencia y desarrollo de la vida pudo ser solo otra consecuencia de las ecuaciones de la física de partículas. Incluso si nadie tiene todavía una idea

sobre los detalles sobre la que se pueda trabajar, es posible especular que la aparición de la vida fue el producto de procesos químicos regidos por las leyes de la física y que la evolución posterior se debe igualmente a mutaciones químicas y a la selección natural que son solo consecuencias hipercomplejas de los principios físicos. Incluso si queda el problema residual de cómo explicar la conciencia en términos físicos, el punto de vista ortodoxo naturalista es que la biología se explica en principio completamente por la física y la química y que la psicología evolutiva proporciona los rasgos generales de cómo todo lo que distingue la vida humana puede ser considerado también como una consecuencia extremadamente compleja del comportamiento de partículas físicas de acuerdo con ciertas leyes fundamentales. Esto, por último, incluirá una explicación de las capacidades cognoscitivas que nos capacitan para descubrir dichas leyes.

Encuentro incomprensible que este punto de vista sobre las cosas deba ser considerado, más o menos, como evidente por sí mismo, como creo que lo es comúnmente. Todo el mundo reconoce que hay cantidades ingentes de cosas que no conocemos y que ante nosotros se extienden enormes posibilidades para progresar en su comprensión. Pero el científico naturalista reclama saber cuál será la forma de ese progreso y saber en particular que la inteligibilidad mentalista, teleológica o valorativa han sido superadas por mejores y más fundamentales formas de comprensión. Se asume no solo que el orden natural es inteligible sino que su inteligibilidad tiene cierta forma que se funda en las leyes físicas más simples y más unificadas que gobiernan los más simples y menores elementos. Y de eso se sigue todo lo demás. Eso es lo que el científico optimista quiere decir con teoría del todo. En la medida en que las leyes básicas no son ellas mismas verdades necesarias, permanece el problema de por qué esas leyes funcionan. Pero quizás parte de la llamada de atención de esta concepción es que si las leyes son bastante simples podemos descansar en ellas y alegrarnos de decir que así es exactamente como son las cosas. Después de todo, ¿cuál es la alternativa?

Y esa es realmente mi pregunta. La implausibilidad del programa reductivo que se necesita para defender la completitud de esta clase de

naturalismo da motivos para intentar buscar alternativas, alternativas que convierten a la mente, al significado y al valor en algo tan fundamental como la materia y el espacio-tiempo a la hora de explicar lo que hay. Los elementos fundamentales y las leyes de la física y la química han sido descubiertos para explicar el comportamiento del mundo inanimado. Se necesita algo más para explicar cómo puede haber criaturas pensantes y conscientes cuyos cuerpos y cerebros están compuestos de esos elementos. Si queremos intentar comprender el mundo como un todo debemos comenzar con un espectro adecuado de datos y entre esos datos se deben incluir los hechos más evidentes sobre nosotros mismos.

4

Para marcar los límites del terreno en el que debemos proceder a la búsqueda de tal comprensión, me gustaría decir algo sobre el polo opuesto al materialismo, es decir, la postura que afirma que es la mente, más que la ley física, la que aporta el nivel fundamental de explicación de todo, incluyendo la explicación de las mismas leyes físicas básicas y universales. Este punto de vista es comúnmente conocido como teísmo y es, a su manera, una explicación de la existencia y contenido del mundo natural. Es la manera más directa de invertir el orden materialista de explicación que justifica la mente como consecuencia de la ley física. En vez de ello, el teísmo convierte a la ley física en consecuencia de la mente.

Como respuesta a la exigencia de una forma omniabarcante de comprensión, el teísmo interpreta la inteligibilidad en términos de intención o propósito y se opone a un punto final en la explicación que sea puramente descriptivo. En los límites exteriores del mundo, que abarca todo en él, inclusive el orden natural gobernado por leyes establecido por la ciencia, el teísmo sitúa alguna clase de mente o intención responsable de la naturaleza tanto física como mental del universo. En la medida en que la mente divina tiene que ser postulada como parada en la búsqueda de la comprensión deja incompleto el proceso, tanto como lo hace la explicación del materialismo puramente descriptivo.

Para que la explicación materialista o la teísta proporcionaran una comprensión completa del mundo debería ocurrir que o las leyes de la física o la existencia y propiedades de Dios y, por tanto, de su creación, no pudieran ser otras que las que son. Los físicos no creen generalmente lo primero[12], los teístas tienden a creer lo segundo. Eso no significa que una concepción del mundo teísta deba ser determinista: la naturaleza y esencia de Dios puede llevarle a crear leyes probabilistas y seres con voluntad cuyas acciones son elecciones libres. Pero algún tipo de intención divina está en la base de la explicación de la totalidad.

Lo interesante del teísmo, incluso para un ateo, es que intenta explicar de otra forma lo que no parece capaz de ser explicado por la ciencia física. Las deficiencias de la imagen del mundo naturalista y reduccionista me parecen reales. Hay cosas que la ciencia tal y como la concebimos en el presente no nos ayuda a entender y, que podamos ver desde la estructura interna de la ciencia física, no va a explicar. Parecen reclamar una forma de comprensión más fuertemente mentalista o incluso normativa. El teísmo sostiene esa conclusión al atribuir los fenómenos mentales que aparecen dentro del mundo a la obra de una fuente mental comprehensiva de la cual esos fenómenos son versiones en miniatura.

Sin embargo el teísmo no es más creíble que el materialismo como concepción del mundo comprehensiva. Lo que me interesa particularmente está en el terreno que se encuentra entre ellos. Creo que estas dos concepciones radicalmente opuestas sobre la inteligibilidad última no pueden agotar las posibilidades. Todas las explicaciones tienen que concluir de alguna forma. Tanto el teísmo como el materialismo dicen que en el último nivel hay una forma de comprensión. ¿Pero sería posible una concepción laica alternativa que reconociera a la mente y a todo lo que implica no como la expresión de una intención divina sino como un principio fundamental de la naturaleza al mismo nivel que la ley física? ¿Podría tomar la forma de una concepción unificada del orden natural incluso si intenta albergar un conjunto más rico de materiales que los elementos de la física matemática? Pero permítaseme primero decir un poco más, por fines dialécticos, sobre la oposición entre teísmo y naturalismo materialista y lo que le falta a cada uno de ellos.

5

La mayor diferencia entre las formas de inteligibilidad que hemos considerado se da en la comprensión de nosotros mismos. Ese es también el marco de las más acaloradas batallas sobre lo que la ciencia física puede y no puede explicar. Tanto teísmo como naturalismo evolucionista son intentos de comprendernos desde fuera usando fuentes muy diferentes. El teísmo ofrece una comprensión vicaria asignando el rol principal a una mente trascendente cuyos propósitos y comprensión del mundo no podemos compartir del todo pero que hace posible creer que el mundo es inteligible, incluso si no lo es para nosotros. La forma de esta comprensión trascendente se genera por extrapolación de la auto-comprensión psicológica natural que tenemos de nuestras propias intenciones. Por el contrario, el naturalismo evolucionista extrapola todo, incluyendo nosotros mismos, a una forma de comprensión científica que hemos desarrollado para aplicarla a otras partes del mundo. Estos dos acercamientos comparten la ambición de dar razón de nosotros mismos desde una comprensión que nace del intento de trascender nuestros puntos de vista subjetivos; eso es tan importante como la diferencia entre ellas.

¿Qué es lo que justifica, si lo hay, este común deseo de trascendencia? ¿No es suficiente intentar comprendernos solo interiormente, lo que ya es bastante difícil? Pero ese deseo es irresistible, como si no pudiéramos continuar legítimamente en la vida solo desde el punto de vista que naturalmente ocupamos en el mundo sino que necesitáramos trascendernos en una concepción del mundo más amplia. Y para valer, esa concepción del mundo más amplia debe bastarse por sí misma.

Cualquier comprensión del mundo externo, aunque sea trascendente, comienza con nuestro punto de vista (¿cómo podría ser de otra forma?) y se supone que es consistente con las líneas principales de ese punto de vista incluso si también proporciona una crítica significativa y una revisión o una ampliación del propio punto de vista. Respecto al conocimiento humano, por ejemplo, tanto el teísmo como el naturalismo intentan explicar cómo podemos confiar en nuestras facultades para comprender el mundo que nos rodea. En un extremo está la validación teísta de Descartes de la percepción

y el razonamiento científico por la prueba de que Dios, que es responsable de nuestras facultades, no puede engañarnos. En el otro extremo está la epistemología naturalizada que argumenta que se puede confiar en que las facultades perceptivas y cognoscitivas desarrolladas por selección natural nos conducen a creencias verdaderas.

Ninguna de estas propuestas proporciona una defensa contra el escepticismo radical, es decir, contra la posibilidad de que nuestras creencias sobre el mundo sean falsas. Tal defensa sería inevitablemente circular en tanto que la confianza que podamos tener en la verdad, ya sea de una explicación teísta o evolucionista de nuestras capacidades cognoscitivas, tendría que depender del ejercicio de esas capacidades. Para el teísmo, este es el famoso círculo cartesiano; pero hay un círculo naturalista análogo[13]. Pero además de eso, el naturalismo evolucionista ofrece una explicación de nuestro conocimiento que es muy inadecuada cuando se aplica a las capacidades de generar conocimiento que creemos tener. Volveré a esa cuestión más adelante.

Aunque estos dos proyectos de auto-comprensión no refuten el escepticismo, creo que hay una legítima aspiración de trascendencia que es más modesta y, quizás, más realista. Puede que no seamos capaces de descartar la posibilidad escéptica y puede que no seamos capaces de fundar nuestra capacidad habitual de comprensión en algo en lo que podamos tener una mayor confianza; pero cabría la posibilidad de mostrar que podemos mantener una natural confianza en el ejercicio del entendimiento a pesar de lo contingente de nuestra naturaleza y de nuestra génesis. La esperanza no está en descubrir un fundamento que convierta nuestro conocimiento en irrefutable, sino en encontrar una forma de comprendernos a nosotros mismos que no se auto-destruya radicalmente y que no requiera negar lo obvio. El objetivo sería ofrecer una imagen plausible de cómo nos adaptamos al mundo.

6

Pero incluso en esta empresa más modesta tanto el teísmo como el reduccionismo naturalista se quedan cortos. El teísmo no ofrece una

explicación suficientemente sustantiva de nuestras capacidades y el naturalismo no ofrece una suficientemente buena.

Una explicación teísta tiene la ventaja, sobre otra naturalista reductiva, que admite la realidad de más de lo que resulta evidente e intenta explicarlo todo. Pero incluso si el teísmo se llena con las doctrinas de una religión particular (que no será accesible a la evidencia y a la sola razón), ofrece una explicación muy parcial de nuestro lugar en el mundo. Expresa la hipótesis de que la explicación de más alto orden de cómo las cosas encajan unas con otras es de un cierto tipo, a saber, intencional o con un propósito, pero no tiene nada más que decir sobre cómo obra esa intención excepto que se encuentra entre los asuntos a explicar.

La idea no es una simpleza porque cualquier explicación intencional envuelve alguna creencia interpretativa, incluso sobre Dios. Se debe pensar en un agente intencional como en alguien que tiene objetivos que ve como buenos, así que los objetivos no pueden ser arbitrarios; una explicación teísta comportará inevitablemente alguna idea del valor y una religión en particular puede concretar eso mucho más aunque también le surja el famoso problema del mal. En lo que a mí respecta, además de la dificultad de creer en Dios, veo que la desventaja del teísmo como respuesta al deseo de comprensión abarcadora no es que no ofrezca explicaciones sino que no lo hace en la forma de una explicación comprehensiva del orden natural. El teísmo compele a una búsqueda de la inteligibilidad fuera del mundo. Si Dios existe no es parte del orden natural sino un agente libre no regido por leyes naturales. Puede actuar en parte creando un orden natural, pero lo que hace directamente no puede ser parte de ese orden.

Una comprensión teísta, para los que encuentran convincente entender el mundo como expresión de una intención divina, dejaría intacta nuestra confianza natural en nuestras facultades cognoscitivas. Pero no sería la clase de comprensión que explica *cómo* seres como nosotros se adaptan al mundo. La clase de inteligibilidad que se estaría echando en falta es la inteligibilidad del orden natural mismo, una inteligibilidad desde dentro. Esa clase de inteligibilidad puede ser compatible con algunas formas de teísmo: si Dios creara un orden natural autónomo que luego deja intacto. Pero no es compatible con una explicación teísta directa de ciertos

acontecimientos del mundo ya que, de otra forma, parecerían hechos brutos, tal como la creación de la vida de la materia inerte o el nacimiento de la conciencia o de la razón. Tales hipótesis intervencionistas equivalen a una negación de que hay un orden natural comprehensivo. Están en parte motivadas por una creencia que considero correcta, a saber, que hay poca o ninguna posibilidad de que estos hechos dependan de nada más que de las leyes de la física. Pero otra respuesta alternativa sería pensar que puede haber un tipo de explicación sistemática completamente diferente de la naturaleza, una que no los convierta ni en hechos brutos que están más allá de la explicación ni en productos de la intervención divina. Esa, sin lugar a dudas, es mi infundada preferencia intelectual.

El problema de las teorías naturalistas es diferente: mejor que ser seguras pero insuficientemente explicativas, las teorías materialistas intentan hacer internamente inteligible el orden natural explicando nuestro lugar en él sin referencia a algo externo. Pero las explicaciones que proponen no son suficientemente buenas. El naturalismo evolucionista proporciona una explicación de nuestras capacidades que mina nuestra confianza en ellas y, al hacerlo, se destruye a sí mismo. Tendré más que decir sobre estos problemas del reduccionismo más tarde; permítaseme aquí esbozarlos brevemente.

Inevitablemente, cuando construimos una comprensión naturalista, estamos confiando, por una parte, en nuestras capacidades de «dar-sentido» para crear un sistema que dará sentido al resto. Confiamos en la teoría evolucionista para analizar y evaluar todo, desde nuestro conocimiento lógico y probabilístico hasta nuestro sentido moral. Esto refleja el punto de vista de que la ciencia empírica es la única forma segura y privilegiada de comprensión y que podemos confiar en otras formas solo en la medida en que puedan ser validadas por una explicación científica de cómo y por qué funcionan. Eso requiere confianza en alguna de nuestras propias facultades. Pero se piensa que algunas facultades merecen más confianza que otras, e incluso si no pudiéramos proporcionarles una justificación no circular, debemos creer al menos que no son destruidas por la explicación externa que se propone de sus fuentes y operaciones. Debe permanecer intacto un

centro de confianza cognoscitiva incluso si algunas otras facultades se tornan dudosas debido a su origen evolutivo.

En su estructura eso es todavía el ideal cartesiano, pero con la función directora desempeñada por la teoría evolucionista en lugar de por una demostración *a priori* de la benevolencia divina. Pero estoy de acuerdo con Alvin Plantinga en que, a diferencia de la benevolencia divina, la aplicación de la teoría evolucionista a la comprensión de nuestras propias capacidades cognoscitivas mina, aunque no necesariamente destruye completamente, nuestra confianza en ellas[14]. Los mecanismos de formación de creencias que tienen ventaja selectiva en la lucha diaria por la existencia no garantizan nuestra confianza en la construcción de explicaciones teóricas del mundo como totalidad. Creo que la hipótesis evolucionista implicaría que aunque se *pudiera* confiar en nuestras capacidades cognoscitivas no tenemos en realidad razones para confiar en ellas como las tenemos de ordinario cuando las usamos directamente, tal y como hacemos en ciencia. En particular, no se explica por qué se justifica la confianza en ellas para corregir otras disposiciones cognoscitivas que pueden extraviarnos, aunque puedan ser igualmente naturales e igualmente susceptibles de explicación evolutiva. La historia evolucionista deja la autoridad de la razón en una posición mucho más débil. Eso es incluso más claro respecto de nuestras capacidades morales y de otras capacidades normativas, en las que frecuentemente confiamos para corregir nuestros instintos. Estoy de acuerdo con Sharon Street en que una comprensión evolucionista requeriría, casi con certeza, abandonar el realismo moral, la convicción natural de que nuestros juicios morales son verdaderos o falsos independientemente de nuestras creencias[15]. El naturalismo evolucionista implica que no debemos tomar en serio ninguna de nuestras convicciones, incluyendo la imagen científica del mundo sobre la que depende el mismo naturalismo evolucionista.

Defenderé esas afirmaciones en los capítulos siguientes, pero permítanme decir aquí qué pasaría si son correctas. El fracaso del naturalismo evolucionista para proporcionar una forma de comprensión trascendente que no destruya nuestra confianza en nuestras facultades naturales no debe llevarnos a abandonar la búsqueda de una comprensión

trascendente. No hay motivo para otorgar nuestra confianza a la verdad objetiva de nuestras creencias morales o, por lo mismo, nuestra confianza a la verdad objetiva de nuestro razonamiento matemático o científico, si la hacemos depender de si es consistente con la suposición de que estas capacidades son el producto de la selección natural. Dado cómo son las explicaciones evolucionistas especulativas de las facultades mentales humanas parecen un fundamento demasiado débil para poner en cuestión las formas más básicas de pensamiento. Nuestra confianza en la verdad de las proposiciones que parecen evidentes a la reflexión no debería tambalearse tan fácilmente (y, añadiría, no pueden tambalearse por esta clase de fundamentos sin algún tipo de falsa conciencia).

Parece razonable aplicar el test igualmente en la dirección opuesta, a saber, evaluar la hipótesis sobre el universo y sobre cómo hemos venido a la existencia refiriéndonos a los juicios ordinarios en los que tenemos una altísima confianza. Es razonable creer que la verdad sobre qué tipo de seres somos y cómo el universo nos produjo es compatible con tal confianza. Después de todo, todo lo que creemos, incluso las teorías cosmológicas más lejos de nuestro alcance, tiene que fundarse por último en el sentido común y sobre lo que es simplemente innegable. La prioridad explicativa dada al naturalismo evolucionista, a pesar de sus nada plausibles conclusiones sobre otros asuntos, es debido, pienso, al consenso laico de que esta es la única forma de comprensión externa de nosotros mismos que proporciona una alternativa al teísmo, que es rechazado sin evidencia alguna como una mera proyección de la función que nos auto-otorgamos en el universo.

7

Incluso si ni el naturalismo evolucionista ni el teísmo proporcionasen el tipo de comprensión abarcadora que estamos buscando, no se debería considerar amenazada nuestra confianza en la operación de nuestra razón aunque su presencia en el mundo siguiera siendo un misterio. Podemos continuar a la espera de una comprensión trascendente que no sea ni teísta ni reduccionista. Pero esto significa rechazar también una tercera alternativa al problema que para mí no resulta sostenible aunque tiene seguidores

distinguidos, a saber, abandonar el proyecto de la comprensión externa y en su lugar limitarnos a la por sí misma formidable tarea de comprender nuestro punto de vista del mundo desde dentro. La ciencia física es un aspecto del punto de vista humano, pero puede coexistir con los otros aspectos sin subsumirlos. Este método pluralista es lo que P. F. Strawson llama «metafísica descriptiva»[16] y tiene mucho en común con la concepción antimetafísica de Wittgenstein de la tarea propia de la filosofía.

Pero así como la comprensión interna es ciertamente valiosa y una condición previa esencial de un proyecto más trascendente, no veo cómo podemos pararnos ahí y no buscar también una concepción externa de nosotros mismos. Para abstenernos tendríamos que creer que la búsqueda de una única realidad es una ilusión porque hay muchas clases de verdad y muchas clases de pensamiento expresados en muchas formas diferentes de lenguaje que no pueden ser armonizados por una concepción de un mundo única en la que se fundamente toda verdad. Esa es una opción tan radical como cualquiera de sus alternativas[17].

El problema sigue planteado, tanto si lo respondemos como si no. Incluso si concluyéramos que la explicación materialista de nosotros mismos es incompleta —incluyendo su desarrollo por medio de la teoría evolucionista— aún resta el asunto de que somos productos de una larga historia del universo desde el *big bang* y que descendemos de las bacterias tras miles de millones de años de selección natural. Esa es parte de la verdadera comprensión externa de nosotros mismos. La pregunta es cómo podemos combinar esa explicación con las demás cosas que conocemos —incluyendo las formas de razón sobre las que esa misma conclusión se funda— en una concepción del mundo que no se destruya a sí misma.

Nuestra propia existencia nos enfrenta con el hecho de que el mundo genera seres conscientes capaces de reconocer razones para la acción y la creencia, distinguir algunas verdades necesarias y evaluar la evidencia de hipótesis alternativas sobre el orden natural. No sabemos cómo ocurre esto, pero es difícil no creer que haya alguna explicación de tipo sistemático, una explicación ampliada del orden del mundo.

Si encontramos innegable, como deberíamos, que nuestros razonamientos lógicos y morales más claros son objetivamente válidos,

estamos en el primer peldaño de la escalera. No nos compromete a ninguna interpretación particular de lo normativo, pero creo que se requiere algo más. No podemos mantener ese tipo de resistencia a cualquier explicación ulterior que se llama quietismo. La confianza que sentimos dentro sobre nuestro propio punto de vista demanda ser completada por un punto de vista más comprehensivo de nuestra inclusión en el mundo. Mientras tanto continuamos usando la percepción y la razón para construir teorías científicas del mundo natural, aunque no tengamos una explicación externa convincente de por qué existen esas facultades que sea consistente con nuestra confianza en su fiabilidad: ni una explicación naturalista ni una teísta cartesiana. La existencia de mentes conscientes y su acceso a las verdades evidentes de la ética y la matemática están entre los datos que una teoría del mundo y de nuestro lugar en él tiene aún que explicar. Son claramente parte del asunto, tanto como los datos sobre el mundo físico proporcionados por la percepción y las conclusiones del razonamiento científico sobre lo que mejor explicaría esos datos. No podemos sin más suponer que la categoría más tardía de pensamiento tiene prioridad sobre las otras de tal manera que lo que no puede explicar no sea real.

Para que una forma adecuada de comprensión fuera una alternativa al materialismo, tendría que incluir elementos mentales y racionales de algún tipo. Pero mi posición es que podrían pertenecer al mundo natural y no necesitan implicar una mente individual trascendente y mucho menos un ser perfecto. El hecho inexorable que tiene que incluirse en una concepción completa del universo es que la aparición de organismos vivos ha dado lugar por último a la conciencia, la percepción, el deseo, la acción y la formación tanto de creencias como de intenciones sobre bases racionales. Si todo esto tiene una explicación natural las posibilidades serían inherentes al universo mucho antes de que hubiera vida e inherentes en la vida temprana muchísimo antes de la aparición de los animales. Una explicación satisfactoria mostraría que la realización de estas posibilidades no era algo desvaídamente improbable, sino una posibilidad significativa dadas las leyes de la naturaleza y la composición del universo. Revelarían a la mente y la razón como aspectos básicos de un orden natural no materialista.

Esto no es triunfalismo antropocéntrico. Todo el reino animal, la interminable generación de insectos y arañas en sus enormes y extravagantes poblaciones, plantean esta misma pregunta sobre el orden de la naturaleza. No hemos observado vida en ningún lugar más que sobre la Tierra, pero ningún hecho natural es cosmológicamente más significativo. Sin embargo, por mucho que lleguemos a entender, como estamos en proceso de hacer, las bases químicas de la vida y de su evolución, el fenómeno todavía requiere fundamentos enormemente ampliados para ser inteligible.

Resumiendo: las deficiencias del materialismo y del teísmo como concepciones trascendentes y la imposibilidad de abandonar la búsqueda de un punto de vista trascendente de nuestro lugar en el universo nos conduce a la esperanza de una comprensión ampliada pero todavía naturalista que evite el reduccionismo psicofísico. La naturaleza esencial de tal comprensión no consistiría en explicar la aparición de la vida, la conciencia, la razón y el conocimiento como efectos colaterales accidentales de las leyes físicas de la naturaleza ni como el resultado de una intervención intencional en la naturaleza sino como una consecuencia nada sorprendente, si no inevitable, del orden interno que gobierna el mundo natural. Ese orden tendría que incluir la ley física, pero si la vida no es solo un fenómeno físico, el origen y evolución de la vida y la mente no resultará explicable solo por la física y la química. Se necesitará una forma de explicación ampliada, pero unificada, y sospecho que tendrá que incluir elementos teleológicos.

Todo lo que puede hacerse en este escenario de la historia de la ciencia es argumentar para que se reconozca el problema, no ofrecer soluciones. Pero quiero, con más detalle, empezar a hacer más pequeños algunos de los obstáculos y sus consecuencias comenzando con el caso más claro.

CAPÍTULO 2

LA CONCIENCIA

1

La conciencia es el obstáculo más claro para un naturalismo comprehensivo que pretenda descansar solo sobre las bases de la ciencia física. La existencia de la conciencia parece implicar que la descripción física del universo, a pesar de su riqueza y poder explicativo, es solamente una parte de la verdad y que el orden natural está lejos de ser tan simple como lo sería si la física y la química lo explicaran todo. Si tomamos este problema seriamente y desarrollamos sus implicaciones, amenaza con deshacer por completo la imagen naturalista del mundo. Pero es muy difícil imaginar alternativas viables.

Permítaseme comenzar con una breve historia de lo que nos ha conducido a la situación presente. El problema moderno mente-cuerpo surgió de la revolución científica del siglo XVII como resultado directo del concepto de realidad física objetiva que dirigió esa revolución. Galileo y Descartes hicieron una escisión conceptual crucial al proponer que la ciencia física debería proporcionar una descripción cuantitativa matemáticamente precisa de una realidad externa desplegada en el espacio y en el tiempo, una descripción que se limita a cualidades primarias espacio-temporales tales como forma, tamaño y movimiento y a las leyes que gobiernan las relaciones entre ellas. Las apariencias subjetivas, por otra parte, —cómo aparece este mundo físico a la percepción humana— se atribuyeron a la mente, y las cualidades secundarias —el color, el sonido y el olor— se analizaron relacionalmente, en términos del poder de las cosas

físicas que actúan sobre los sentidos para producir esas apariencias en las mentes de los observadores. Era esencial excluir o sustraer las apariencias subjetivas y la mente humana —junto con las intenciones y propósitos humanos— del mundo físico para permitir desarrollar esta poderosa pero reductiva concepción espacio-temporal de la realidad física objetiva.

Sin embargo, la exclusión de todo lo mental del ámbito de la moderna ciencia física estaba destinada a ser finalmente cuestionada. Nosotros los humanos somos partes del mundo y el deseo de una imagen unificada del mundo es irreprimible. Resulta natural buscar esa unidad ampliando el alcance de la física y la química a la luz de sus grandes éxitos en la explicación de tanto del orden natural. Estos éxitos han adquirido la forma de una reducción seguida por una reconstrucción: descubriendo los elementos básicos de los que todo se compone y mostrando cómo se combinan para producir la complejidad que observamos.

Hoy resulta claro que nuestros cuerpos y sistemas nerviosos centrales son partes del mundo físico, compuestos de los mismos elementos que todo lo demás y completamente descriptibles en los términos de las versiones modernas de las cualidades primarias, más sofisticadas pero todavía definidas matemática y espacio-temporalmente. La biología molecular ha incrementado el conocimiento de nuestra propia composición física, funcionamiento y desarrollo. Finalmente, hasta donde podemos saber, nuestras vidas mentales, incluyendo las experiencias subjetivas y las de otras criaturas, están fuertemente conectadas con —y son probablemente dependientes de— eventos físicos en nuestros cerebros y de la interacción física de nuestros cuerpos con el resto del mundo físico.

Quizás sea que estos desarrollos en neurofisiología y biología molecular han fomentado la esperanza de incluir la mente en una concepción física única del mundo; de todas formas, el consenso en esa dirección es reciente. Descartes pensó que no podía hacerse, que la mente y la materia son ambas completamente reales e irreductiblemente distintas aunque interactúen. Desde el punto de vista dualista, la ciencia física se define por excluir lo mental como tema de su disciplina. Ha habido siempre resistencia al dualismo, pero durante varios siglos después de Descartes, su expresión más destacada fue el idealismo, el punto de vista de que la mente es la

realidad última y que el mundo físico es de alguna manera reductible a ella. Este intento de superar la división desde lo mental va desde Berkeley —que rechazó la distinción entre cualidad primaria y secundaria y sostuvo que las cosas físicas son ideas en la mente de Dios— al positivismo lógico, que analizó el mundo físico como una construcción a partir de los datos de los sentidos. Después, en un inesperado giro histórico cuyas causas son obscuras, el idealismo fue mayoritariamente sustituido en la filosofía analítica de la última parte del siglo XX por intentos de unificación en la dirección opuesta, es decir, que partían de lo físico.

El materialismo es el punto de vista que sostiene que solamente el mundo físico es irreductiblemente real y que la mente, si es que existe tal cosa, debe encontrar su cabida en él. Esta posición continuaría el progreso de la ciencia física, a través de la biología molecular, para clausurarla del todo incorporando la mente a la realidad física objetiva de la cual fue inicialmente excluida. Lo que esta postura cree es que la física carece de problemas filosóficos y el objetivo principal al que se opone es la imagen dualista de Descartes del fantasma en la máquina. Su tarea es ofrecer una alternativa y aquí comienzan una serie de fracasos.

Una estrategia para incluir lo mental en la imagen física del mundo es el conductismo conceptual, ofrecido como análisis de la naturaleza real de los conceptos mentales. Se intentó en varias versiones. Los fenómenos mentales se identificaron según sus diferentes variedades con la conducta, con disposiciones conductuales o con formas de organización conductual. En otra versión, asociada a Ryle e inspirada por Wittgenstein[18], los fenómenos mentales no fueron identificados con nada, tanto físico como no físico; se dice que los nombres de los estados y procesos mentales no están referidos a expresiones. En vez de eso, los conceptos mentales se explicaron en términos de sus condiciones conductuales observables de aplicación, como son criterios conductistas o condiciones de asertabilidad, antes que según condiciones de verdad conductual.

Todas estas estrategias son esencialmente verificacionistas. Suponen que todo lo que se necesita decir sobre el contenido de un estado mental es lo que verificaría o confirmaría o justificaría su aserción desde el punto de vista de un observador. De una forma u otra reducen los atributos mentales

a las condiciones de lo externamente observable con el criterio de que atribuimos estados mentales a los otros. Si fuera así, esto situaría de forma obvia a la mente cómodamente en el mundo físico.

Es verdad que los fenómenos mentales tienen manifestaciones conductuales que nos suministran la evidencia principal para afirmarlos de otras criaturas. Pero todas estas teorías resultan insuficientes como análisis de lo mental porque omiten algo esencial que descansa más allá de los fundamentos externamente observables para atribuir estados mentales a otros, a saber, el aspecto del fenómeno mental que es evidente desde la primera persona, el punto de vista interior del sujeto consciente: por ejemplo, la forma en que el azúcar te sabe o la forma en que ves el rojo o sientes ira, cada una de las cuales parece ser algo más que las respuestas conductuales y las capacidades discriminatorias que explican estas experiencias. El conductismo omite el estado mental interior mismo.

En los años 1950 se propuso una alternativa, un camino no analítico al materialismo, uno que en un sentido reconocía que lo mental es algo dentro de nosotros cuya conducta observable exteriormente es una mera manifestación. Fue la teoría de la identidad psico-física presentada por U. T. Place y J. J. C. Smart[19] no como análisis conceptual sino como una hipótesis científica. Sostiene que los eventos mentales son eventos físicos en el cerebro: $\psi = \phi$ (donde ψ es un evento mental como el dolor o una sensación del sentido del gusto y ϕ es el correspondiente evento físico en el sistema nervioso central). En tanto que no es una verdad conceptual no puede ser conocida *a priori;* se supone que es una identidad teórica como «Agua = H_2O» y solo puede ser confirmada por el futuro desarrollo de la ciencia.

El problema es que esta identidad no analítica despierta una pregunta ulterior: ¿Qué es ϕ que lo hace también ψ? Debe ser alguna propiedad conceptualmente distinta de las propiedades físicas que definen ϕ. *¿Qué se requiere para que la identidad sea una verdad científica y no una conceptual?*[20]. Claramente los materialistas no querrán dar una respuesta dualista, es decir, que ϕ es ψ porque tiene una propiedad no física además de las suyas fisiológicas (v. g., una cualidad empírica no física). Pero tienen que dar alguna respuesta y tiene que ser una respuesta consistente con el

materialismo. Así pues, los defensores de la teoría de la identidad, a pesar de sus deseos de evitar el análisis conceptual, tienden a volver a diferentes tipos de conductismo analítico para estudiar el carácter mental de los procesos cerebrales de una forma que evite el dualismo. Lo que hace del proceso cerebral un proceso mental, sugieren, no es una propiedad adicional intrínseca sino una relacional, una relación a la conducta física.

Un elemento causal se añadió entonces al análisis: «el estado interior que típicamente causa cierta conducta y es causado por ciertos estímulos». Esto surgió por la necesidad de explicar las dos referencias a un mismo objeto, y que no son sinónimas, que ocurren en un estado de identidad no conceptual. El materialista tiene que explicar cómo «dolor» y «estado cerebral» se pueden referir a lo mismo incluso cuando su significado no es el mismo y explicarlo sin apelar a algo no físico para dar razón de la referencia de «dolor»[21].

Estas estrategias han ido tomando formas cada vez más sofisticadas bajo los nombres de conductismo causal, funcionalismo y otras teorías de cómo los conceptos mentales pueden referirse a estados del cerebro en virtud de la función causal de estos estados para controlar la interacción entre el organismo y su medio ambiente. Pero todas estas estrategias son insatisfactorias debido a la misma vieja razón: incluso con el cerebro añadido a su imagen del mundo, omiten algo esencial sin lo cual no habría mente. Y lo que omiten es exactamente lo que deliberadamente fue excluido del mundo físico por Descartes y Galileo para formar el concepto moderno de lo físico, a saber, las apariencias subjetivas.

Saúl Kripke señaló otro problema. Los teóricos de la identidad toman como modelo de $\psi = \phi$ otras identidades teóricas como «Agua = H_2O» o «Calor = Movimiento molecular». Pero Kripke argumenta que estas identidades son verdades necesarias (aunque no conceptuales y no *a priori)* mientras que la relación ψ/ϕ es contingente[22]. Este fue el fundamento del argumento de Descartes para el dualismo. Descartes dijo que en tanto que podemos concebir la existencia de la mente sin el cuerpo físico, y viceversa, no pueden ser una sola cosa[23].

Considérese «Agua = H_2O» como una típica identidad teórica descubierta científicamente. Eso significa que el agua no es *nada sino* H_2O.

No puedes tener H_2O sin agua y no necesitas nada más que H_2O para tener agua. Es agua incluso si no hay nadie alrededor para verla, sentirla o gustarla. Ordinariamente identificamos el agua con sus cualidades perceptibles, pero nuestras experiencias perceptivas no son parte del agua; son solo efectos que tiene en nuestros sentidos. Las propiedades intrínsecas del agua, su densidad, conductividad eléctrica, índice de refracción, ser líquida entre 0 y 100 grados centígrados, etc., se explican completamente por H_2O y sus propiedades. Las propiedades físicas del H_2O bastan por ellas mismas para el agua.

Así, si ψ es realmente ϕ en este sentido, y nada más, entonces ϕ por ella misma, una vez que se comprenden sus propiedades físicas, debería igualmente bastar para el gusto del azúcar, el sentimiento de dolor o lo que sea que se supone que sea idéntico con ella. Pero no parece serlo. Parece concebible para cualquier ϕ que hubiera un ϕ sin ninguna experiencia en absoluto. Parece que la experiencia del gusto es algo extra, relacionado contingentemente con el estado cerebral, algo *producido* más bien que constituido por el estado cerebral. Así pues, no puede ser idéntico al estado cerebral en la forma en que el agua es idéntica a H_2O.

He dado solo una breve muestra del terreno. Alrededor de estos problemas ha crecido una voluminosa e intrincada bibliografía pero que sirve principalmente para confirmar lo intratables que son. Las múltiples vías muertas en la progresión del materialismo sugieren que puede ser más difícil librarse del dualismo ψ/ϕ introducido en el nacimiento de la ciencia moderna de lo que mucha gente imagina. Incluso ha llevado a algunos filósofos al materialismo eliminativo, es decir, a sugerir que los eventos mentales, como los fantasmas y Santa Claus, no existen en absoluto[24]. Pero si no consideramos esa postura como una opción y todavía buscamos una imagen unificada del mundo creo que tendremos que dejar atrás el materialismo. Los sujetos conscientes y sus vidas mentales son componentes de la realidad de los que no podemos escapar y que no son descriptibles por las ciencias físicas.

Sospecho que la apariencia de que la relación entre mente y cerebro es contingente es probablemente una ilusión y que hay una conexión necesaria pero no conceptual que se nos oculta debido a lo inadecuados que son

nuestros conceptos actuales[25]. Los grandes avances científicos requieren a menudo de la creación de nuevos conceptos para postular elementos inobservables de la realidad que se necesitan para explicar cómo las regularidades naturales que inicialmente aparecen como accidentales son de hecho necesarias. La evidencia de que tales cosas existen la da precisamente que, si existen, explicarán lo que de otra forma es incomprensible.

El problema mente-cuerpo es lo suficientemente difícil para que debamos sospechar de los intentos de resolverlo con conceptos y métodos desarrollados para explicar tipos muy diferentes de cosas. En lugar de eso, deberíamos esperar un progreso teórico en este área que requiere una revolución conceptual considerable, al menos tan radical como la teoría de la relatividad, la introducción de los campos electromagnéticos en física, o la misma revolución científica que, a causa de sus intrínsecas restricciones, no puede convertirse en una «teoría de todo» pero que debe verse como un paso en el camino a una forma más general de comprensión. Nosotros mismos somos, a gran escala, ejemplos complejos de algo que es a la vez objetivamente físico desde fuera y subjetivamente mental desde dentro. Quizás los fundamentos de esta identidad penetran el mundo.

2

Hasta ahora he argumentado que las ciencias físicas no nos capacitarán para comprender los centros subjetivos de conciencia, que son irreductibles y constituyen una parte importante del mundo. Pero el fracaso del reduccionismo en filosofía de la mente tiene implicaciones que se extienden más allá del problema mente-cuerpo. El reduccionismo psicofísico es un componente esencial de un programa naturalista más amplio que no puede sobrevivir sin él. Este programa naturalista es metafísico y científico. Sostiene que todo en el mundo es físico y que todo lo que ocurre en el mundo tiene su explicación más básica, tanto si llegamos a saberlo como si no, en la ley física aplicada a las cosas físicas y a sus eventos y sus constitutivos.

Muchos filósofos de la mente, quizás la mayoría, están comprometidos con el proyecto reduccionista; piensan en las dificultades que he descrito simplemente como problemas que necesitan resolverse y que pueden serlo con éxito. Cualquiera que comparta ese punto de vista puede considerar el argumento que sigue como una hipótesis. Su propósito es mostrar que *si* se descarta el reduccionismo psicofísico, toda nuestra comprensión naturalista del universo se verá afectada, no solamente nuestra comprensión de la conciencia. Comenzando con la biología y calando hasta nuestra concepción de los elementos básicos de la realidad, convierte la forma de materialismo estándar actual en algo que no se puede intentar ni siquiera como explicación del mundo físico por la sencilla razón de que nosotros somos parte de ese mundo. Doy por supuesto que esta conclusión hipotética será bienvenida por los reduccionistas porque muestra lo extravagante y costosa que es una posición antirreduccionista en filosofía de la mente.

Los reduccionistas creen que el camino ha sido allanado por lo completa que es la concepción materialista del mundo debido a alguna forma de análisis del rol funcional o causal de lo mental, y que incluye todos los contenidos de conciencia. Una explicación biológica —evolucionista— de la naturaleza y origen de estas capacidades y funciones conductuales por referencia a las cuales la conciencia puede ser analizada proporciona la unión final con la ciencia física más básica. Todo esto envuelve una gran cantidad de especulación y conjetura evolucionista, pero la idea general de cómo la conciencia tiene cabida en el mundo físico queda bastante clara. Se incluye en virtud de la existencia de organismos físicos capaces de ciertos tipos de interacciones conductuales con el mundo, lo cual es a su vez explicado por la variación genética y la selección natural. Si hay un problema en cómo el materialismo pueda explicar la existencia de tales organismos, eso no tiene nada que ver en particular con la conciencia, sino con el problema general de si la teoría evolucionista proporciona realmente las bases para la reducción de la biología a química y a física.

Pero si el programa de análisis de la conciencia en términos de conducta y de sus causas físicas no es viable, surge otro problema. Incluso si la conciencia es algo que no puede ser analizado desde el punto de vista de las propiedades puramente físicas de los organismos, todavía hay que explicar

su aparición como parte de un proyecto más amplio de dar sentido al mundo. Es más, cualquier explicación tal debe dar razón del hecho de que la aparición de la conciencia sobre la Tierra y las diferentes formas que adopta dependen estrechamente del desarrollo evolutivo de aquellas formas físicas de vida que tienen conciencia. No sabemos con precisión qué formas de vida son esas, pero podemos estar razonablemente seguros de que se extienden mucho más allá de nuestra especie. La evolución de la vida debe ser, al menos, parte de la explicación del desarrollo y las formas de la conciencia.

El problema, entonces, es este: ¿qué clase de explicación del desarrollo de estos organismos, incluso una explicación que no excluya a la teoría evolucionista, podría dar razón de la aparición de organismos que no solo están adaptados al medio ambiente sino que son también sujetos conscientes? Brevemente, creo que de eso no puede haber una explicación puramente física. Lo que tiene que explicarse no es solo el encaje de la vida orgánica con un barniz de *qualia* sino la aparición de puntos de vista subjetivos individuales, un tipo de existencia lógicamente distinta de cualquier cosa describible por las solas ciencias físicas. Si la teoría evolucionista es puramente una teoría física, entonces podría en principio proporcionar el marco para una explicación física de la aparición de organismos animales conductualmente complejos con sistema nervioso central. Pero la conciencia subjetiva, si no es reductible a algo físico, no sería parte de esta historia; quedaría completamente inexplicada por la evolución física, incluso si la evolución física de tales organismos es de hecho una condición causalmente necesaria y suficiente para la conciencia.

La mera aserción de tal conexión no es un punto final aceptable. No es una explicación decir solo que el proceso físico de la evolución ha dado como resultado criaturas con ojos, oídos, sistemas nerviosos centrales, y así sucesivamente, y que es simplemente un hecho bruto de la naturaleza que tales criaturas *son* conscientes en las formas en que las conocemos. Identificar una causa no es proporcionar una explicación significativa si no nos ayuda a entender por qué la causa produce el efecto. Lo que defiendo es que, en tanto que el carácter consciente de esos organismos es uno de sus rasgos más importantes, la explicación de la aparición de tales criaturas

debe incluir una explicación de la aparición de la conciencia. No pueden ser problemas separados. Dar razón de su evolución biológica debe explicar la aparición de los organismos conscientes como tales.

Ya que una explicación materialista no puede hacer esto, la versión materialista de la teoría evolucionista no puede ser la verdad completa. A organismos como nosotros no solo les *pasa* que son conscientes; por tanto ninguna explicación del carácter físico de estos organismos puede ser adecuada si no es también una explicación de su carácter mental. En otras palabras, el materialismo es incompleto incluso como teoría del mundo físico ya que el mundo físico incluye organismos conscientes entre sus más conspicuos ocupantes.

Este problema depende solamente de la creencia de que, aunque el reduccionismo es falso, la mente es un fenómeno biológico. En la medida en que lo mental es irreductible a lo físico, la aparición de organismos físicos conscientes queda inexplicada por una justificación naturalista. En una comprensión puramente materialista de la biología, la conciencia tendría que considerarse como un enorme e inexplicable hecho bruto extra del mundo. Si va a ser explicado naturalistamente mediante una comprensión de la vida orgánica, algo fundamental se debe cambiar de nuestra concepción del orden natural que dio lugar a la vida.

¿Qué clase de concepción unificada del mundo natural permitiría la explicación del desarrollo de los organismos vivos, también para explicar el desarrollo de la conciencia? El antirreduccionismo nos permite plantear la cuestión, pero responderla requiere algo más positivo. Y no puede consistir (¡simplemente!) en una revisión de los conceptos fundamentales de la física, aunque sea radical, como ocurrió con la introducción de los campos electromagnéticos o el espacio-tiempo relativista. Si seguimos suponiendo que somos partes del mundo físico y que el proceso evolutivo que nos trajo a la existencia es parte de su historia, entonces se debe añadir algo más a la concepción física del orden natural que nos permita explicar cómo puede dar lugar a organismos que son más que físicos. Los recursos de la ciencia física no son adecuados para este propósito puesto que esos recursos se desarrollaron para dar cuenta de datos de una naturaleza completamente diferente.

La aparición de la conciencia animal es evidentemente el resultado de la evolución biológica, pero este hecho empírico bien confirmado no es todavía una explicación puesto que no proporciona comprensión ni nos capacita para ver por qué iba a esperarse ese resultado o cómo sucedió. En este caso, a diferencia de la aparición de las adaptaciones físicas características de la vida, no es suficiente una explicación por selección natural fundada en la eficacia física para sobrevivir. La selección por eficacia física reproductiva puede haber terminado en la aparición de organismos que son de hecho conscientes y que tienen una variedad observable de diferentes tipos específicos de conciencia, pero no hay explicación física de por qué esto es así, ni otra explicación de ninguna clase que nosotros sepamos.

Para hacer inteligibles hechos de esta clase, una teoría postmaterialista tendría que ofrecer una explicación unificada de cómo las características físicas y mentales de los organismos se desarrollaron juntas y tendría que hacerlo no solo añadiendo una cláusula al uso que dijera que lo mental aparece con lo físico como un extra. La necesidad de una explicación esclarecedora de la aparición de lo mental nos lleva hacia atrás para proponerla como una parte de la comprensión del proceso entero que condujo a ese resultado.

3

Estoy poniendo mucho empeño en la idea de explicación y la meta de inteligibilidad a la que aspira, una meta que supone la inteligibilidad fundamental del universo como se enunció en el capítulo anterior. No todo tiene una explicación en este sentido. Algunas cosas que parecen exigir explicación, como las muertes en tiempo breve de varios parientes cercanos pueden ser solo coincidencias cuyos componentes se explican sin estar relacionados. Pero los esquemas sistemáticos del mundo natural no son coincidencias y no creo que podamos considerarlos como hechos brutos que no requieran explicación. Las regularidades, los patrones y la organización funcional exigen explicación, cuanto más la tienen más frecuentes son. Cuando llegamos a ser conscientes de tales hechos concluimos que hay algo

que no sabemos, algo que, si lo supiéramos, volvería los hechos inteligibles. Y concedo que saber la causa inmediata de algún efecto no siempre lo hace inteligible: que la causa de la conciencia sea la actividad cerebral es un ejemplo excelente.

La explicación, a diferencia de la causalidad, no lo es solo de un evento, sino de un evento bajo una descripción. Una explicación debe mostrar por qué era probable que un evento *de tal tipo* ocurriera. Podemos saber las causas de las muertes cercanas de varios miembros de una familia pero eso no explicará por qué murieron varios miembros de esa familia, a menos que haya alguna relación entre las causas de las muertes individuales que haga probable por anticipado que afecten al grupo, tal y como una venganza o una enfermedad genética.

Otro ejemplo: hay una explicación física de por qué cuando tecleo «3», «+», «5» y «=» en mi calculadora de bolsillo aparece en la pantalla la figura «8». Pero la explicación causal de la forma en la pantalla no es una explicación de por qué el aparato produce la respuesta correcta. Para explicar el resultado bajo esa descripción debemos hacer referencia al algoritmo que rige la calculadora y la intención del diseñador al darle una realización física.

Una ampliación naturalista de la teoría evolucionista que diera razón de la conciencia no se referiría a las intenciones de un diseñador. Pero si aspira a explicar la aparición de la conciencia *como tal*, tendría que ofrecer alguna explicación de por qué la aparición de los organismos conscientes era probable y no simplemente la de los organismos conductualmente complejos.

La explicación de las características puramente físicas de los organismos que hace la teoría evolucionista estándar es bastante compleja, incluso si no se tiene en cuenta la conciencia. Como he dicho antes, la complejidad física y funcional de los resultados impone condiciones muy exigentes a una explicación histórica reduccionista. La teoría de la selección natural, si descansara solamente sobre la intervención de la ley física, tendría que postular que hay una explicación puramente física de por qué no es improbable que las mutaciones accidentales en el material genético hayan generado la gama de variación de fenotipos viables necesarios para

permitir que la selección natural produzca la historia evolutiva que de hecho ha ocurrido sobre la Tierra en los últimos tres mil millones de años[26]. Como cualquier explicación histórica incorporará muchos elementos contingentes, así que la historia particular de la vida no se explicará solo por la teoría evolutiva. Pero las contingencias y sus efectos tienen que ser consistentes con el carácter físico de la teoría. Y para completar el lazo con la física, la explicación tiene que suponer que hay una probabilidad nada desdeñable de que algunas secuencias de pasos, comenzando desde la materia no viva y dependiendo de mecanismos puramente físicos, pudieran al final haber acabado en una molécula replicante capaz de todo esto, que incorpora un preciso código con una longitud de miles de millones de caracteres, junto con los ribosomas que convierten ese código en proteínas[27]. No es suficiente decir: «Algo tenía que pasar, así que ¿por qué no esto?». Encuentro difícil de comprender la confianza que existe entre el *establishment* científico de que todo este escenario dará lugar a una explicación puramente química, excepto como manifestación de un compromiso axiomático con el materialismo reductivo[28].

Pero explicar la conciencia, tanto como la complejidad biológica, como consecuencia del orden natural añade una nueva y completa dimensión de dificultad. Estoy dejando completamente a un lado el dualismo, que abandonaría la esperanza de una explicación integradora. En verdad, el dualismo de la substancia implicaría que la biología no tiene ninguna responsabilidad en absoluto en la existencia de mentes[29]. Lo que me interesa es la hipótesis alternativa de que la evolución biológica es la responsable de la existencia de los fenómenos mentales conscientes, pero que en tanto estos fenómenos no son explicables físicamente, el punto de vista usual de la evolución se debe revisar. No es solo un proceso físico.

Si eso es así, ¿qué se tendría que añadir a la trama física para alcanzar una explicación genuina de la conciencia, una que hiciera la aparición de la conciencia inteligible, y se opusiera a explicar la aparición de ciertos organismos físicos que, simplemente y como una cuestión de hecho, son conscientes? No es suficiente añadir a la explicación física de la evolución la observación posterior de que diferentes tipos de organismos animales, dependiendo de su constitución física, tienen diferentes formas de vida

consciente. Eso presentaría la conciencia de los animales como un misterioso efecto colateral de la historia física de la evolución, que explica solamente los caracteres físicos y funcionales de los organismos.

Elliot Sober me sugirió una vez en esa línea que la conciencia podría ser como el rojo de la sangre, un efecto colateral de estructuras biológicas funcionales que no tiene función por sí mismo ni explicación directa por selección natural. En ese caso la conciencia sería como una enjuta gigante en el sentido de Gould y Lewontin[30] (y una muy afortunada para nosotros). Pero claramente esta simple identificación de una causa no sería una explicación satisfactoria. Sin más, no explicaría por qué los organismos particulares son conscientes ni por qué los organismos conscientes llegan a existir en absoluto.

Para una explicación satisfactoria de la conciencia como tal tendría que urdirse en la trama evolutiva una teoría psicofísica general de la conciencia, una que hiciera inteligible tanto (1) por qué organismos específicos tienen la vida consciente que tienen y (2) por qué los organismos conscientes surgieron en la historia de la vida en la tierra. En este momento tal teoría es una completa fantasía, pero todavía podemos mostrar algunos problemas sobre lo que tendría que conseguir, en particular sobre la relación entre las partes (1) y (2) de la tarea explicativa.

Supóngase que hubiera una teoría general psicofísica que, si pudiéramos descubrirla, permitiera comprender para cualquier tipo de organismo físico por qué tiene o no tiene vida consciente y, si la tuviera, por qué tiene el tipo específico de vida consciente que tiene. Podría llamarse una teoría no-histórica de la conciencia. Cumpliría la tarea (1). Pero creo que incluso si una teoría no-histórica tan potente estuviera unida a una teoría puramente física de cómo estos organismos surgieron a través de la evolución, el resultado no sería una explicación de la aparición de la conciencia como tal. No cumpliría la tarea (2); todavía dejaría la aparición de la conciencia como una concomitancia accidental, y por tanto inexplicada, de algo más, la genuinamente inteligible historia física.

Permítaseme llamar explicación conjuntiva a una en la que *A* explique *B* y *B* tenga como consecuencia *C*. Algunas veces tal conjunción no servirá como explicación de *C* como tal.

Supongamos que *C* es «la muerte de varios miembros de la misma familia», como vimos arriba. Si *A* da la causa independiente de cada una de las cuatro muertes, *B* es la suma de esas muertes, y si son de hecho miembros de la misma familia, entonces *C* es consecuencia de *B* pero no se explica, como tal, por *A*. Podemos explicar por qué murieron cuatro personas que son de hecho miembros de una misma familia sin explicar por qué cuatro miembros de la misma familia murieron.

O considérese la explicación conjuntiva del caso de la calculadora de bolsillo. *A* es la explicación física de lo que pasa cuando tecleo en ella «3+5=», que causa *B,* la aparición en la pantalla de la figura «8». Es un hecho que esta figura es el símbolo del número 8, y las figuras que he tecleado son los símbolos de cierta suma, así que tenemos la consecuencia *C* de que el aparato produjo la respuesta correcta para la suma. Pero, sin más, esto es simplemente una aserción, y no es todavía una explicación de por qué la calculadora me dio una respuesta o la respuesta correcta. Sin el hecho ulterior de que la calculadora fue diseñada para incorporar un algoritmo aritmético y mostrar su resultado en números arábigos, la sola explicación física dejaría el resultado aritmético como algo completamente misterioso. Daría la causa de la figura que apareció en la pantalla, pero no explicaría el número como tal.

Lo correcto parece ser que una explicación conjuntiva que va de *A* a *B* y de *B* a *C* pueda explicar *C* solamente si hay algo ulterior, una relación interna entre la forma en la que *A* explica *B* y la forma en la que *B* explica *C*. En el caso de la familia, esto se satisfaría si la misma enfermedad rara hereditaria matara a las cuatro personas: cada una de ellas desarrolló la enfermedad en parte porque eran miembros de la misma familia. En el caso de la calculadora, la condición se satisface en tanto que el aparato tiene la estructura física y la función que tiene precisamente para incorporar el algoritmo aritmético.

No basta con que *C* debiera ser la consecuencia, incluso la consecuencia necesaria, de *B*, la cual se explica por *A*. Debe haber algo en el mismo *A* que hace a *C* una consecuencia probable. Creo que si *A* es la historia evolutiva, *B* es la aparición de ciertos organismos y *C* son sus conciencias, esto significa que debe aplicarse alguna clase de teoría psicofísica no solo

no-históricamente, al final del proceso, sino también al proceso evolutivo mismo. Ese proceso tendría que ser no solo la historia física de la aparición y desarrollo del los organismos físicos sino también una historia mental de la aparición y desarrollo de los seres conscientes. Y de alguna forma tendría que ser un proceso que haga inteligible los dos aspectos del resultado.

Si, por ejemplo, la explicación de la vida consciente no reductible tuviera que preservar la estructura básica de la teoría evolutiva, implicaría probablemente lo siguiente: (1) Al menos en los escenarios más tardíos, la conciencia desempeña *per se* una función causal esencial en la supervivencia y reproducción de los organismos. (2) Los rasgos de la conciencia que desempeñan esta función se transmiten genéticamente de algún modo. (3) La variación genética entre individuos que aporta los candidatos para la selección natural, al menos hasta cierto punto, es simultáneamente una variación física y mental. (4) Más aún, y de máxima importancia, parece inevitable que estos mecanismos deban ser precedidos por otros en estados más tempranos de la evolución que creen las condiciones de su posibilidad.

Esto significaría abandonar la creencia estándar en que la evolución es dirigida exclusivamente por causas físicas. En verdad, sugiere que la explicación puede tener que ser algo más que física durante todo el tiempo. El rechazo del reduccionismo psicofísico nos deja con un misterio básico sobre el orden natural, siendo uno de los motivos primarios del reduccionismo evitar dicho misterio. Es un doble misterio: primero, la relación entre lo físico y lo mental en cada suceso individual y, segundo, cómo la explicación evolucionista del desarrollo de los organismos físicos puede transformarse en una explicación psicofísica del desarrollo de la conciencia.

La existencia de la conciencia es una de las cosas más comunes del mundo y, al mismo tiempo, de las más sorprendentes. Ninguna concepción del orden natural que no la muestre como algo que era de esperar puede aspirar siquiera a ser un bosquejo de completitud. Y si la ciencia física, sea lo que sea que pueda tener que decir sobre el origen de la vida, deja a la conciencia sin explicar, eso muestra que no puede proporcionar la forma básica de inteligibilidad de este mundo. Debe haber una forma muy

diferente en la que las cosas, tal como son, tengan sentido y eso incluye la forma en la que es el mundo físico. El problema de la mente no puede ser puesto en cuarentena.

4

Dado este vacío en nuestra comprensión, ¿qué clase de explicación tiene sentido imaginar? Hasta aquí he considerado la posibilidad de adiciones o modificaciones a una explicación evolucionista estándar, pero a partir de ahora quiero considerar una escala más amplia de opciones. Todo lo que uno puede hacer es describir posibilidades abstractas, pero para empezar, resulta claro que cualquier explicación tendrá dos elementos: una explicación constitutiva ahistórica de cómo ciertos sistemas físicos complejos son también mentales y una explicación histórica de cómo tales sistemas surgieron en el universo desde sus inicios. Evidentemente la explicación histórica dependerá en parte de la explicación constitutiva correcta en tanto que la última describe el resultado que la primera tiene que explicar. Permítaseme discutir primero las posibilidades constitutivas.

La explicación constitutiva será o reductiva o emergente. Una explicación reductiva explicará el carácter mental de los organismos complejos enteramente en términos de las propiedades de sus constitutivos elementales y, si nos mantenemos en la creencia de que lo mental no puede ser reducido a lo físico, significará que los constitutivos elementales de los que estamos compuestos no son meramente físicos[31]. En tanto que estamos compuestos de los mismos elementos que el resto del universo esta idea tendrá consecuencias amplias y radicales a las que volveré más adelante.

Una explicación emergente, por el contrario, explicará la naturaleza mental de los organismos complejos por principios que unan los estados y procesos mentales con el complejo funcionamiento físico de esos organismos, en particular con sus sistemas nerviosos centrales en el caso de los humanos y de las criaturas que de alguna manera son como ellos. La diferencia con una explicación reductiva es que, mientras que esos principios no reducen lo mental a lo físico, las conexiones que especifican

entre lo mental y lo físico son todas de un orden más alto. Conciernen solamente a organismos complejos y no requieren ningún cambio en la concepción exclusivamente física de los elementos de los que esos organismos están compuestos. Una explicación emergente de lo mental es compatible con una explicación físicamente reduccionista del sistema biológico en el que la mente emerge.

Para cumplir con los requisitos de una explicación genuina de lo mental, una explicación emergente debe ser sistemática. No puede decir solo que cada evento o estado mental sobreviene al estado físico complejo del organismo en el que ocurre. Eso sería la clase de hecho bruto que no constituye una explicación sino que más bien exige una explicación. Pero creo que podemos imaginar una teoría psicofísica de un orden mayor que evitara ver la conexión como un conjunto gigantesco de correlaciones inexplicables y en su lugar comenzara a hacerla inteligible. Los psicólogos fisiólogos están solo comenzando a descubrir la dependencia sistemática de la experiencia visual de los eventos del córtex visual, por ejemplo, pero podemos imaginar que tales indagaciones conducirán a una teoría general. Pero, aunque sea empíricamente exacta, este tipo de teoría del nivel más alto parece insatisfactoria como respuesta final al problema constitutivo. Si la emergencia es toda la verdad, ello implica que los estados mentales están presentes en el organismo como un todo, o en su sistema nervioso central, sin fundamento alguno en los elementos que constituyen el organismo, excepto por el carácter físico de estos elementos que les permite ser ubicados en la forma compleja que, según la teoría del nivel más alto, conecta lo físico con lo mental. Que tales elementos puramente físicos, cuando se combinan de cierta forma, deban necesariamente producir un estado del todo que no está constituido por las propiedades y relaciones de las partes físicas, me parece como si fuera magia, incluso si las dependencias psicofísicas de más alto orden se establecen de forma suficientemente sistemática.

La insatisfacción con un punto final explicativo que relacione estructuras complejas con estructuras complejas es lo que subyace a la constante tendencia hacia el reduccionismo que tiene la ciencia moderna. Es difícil abandonar la creencia de que la verdad de lo complejo debe ser

explicada por la verdad de los elementos. Eso no significa que no puedan emerger nuevos fenómenos en niveles más elevados, pero se espera que sean analizables mediante la naturaleza e interacciones de sus componentes más elementales. Tal inocua emergencia es ilustrada de forma estándar por el ejemplo de la liquidez, la cual depende de las interacciones de las moléculas que componen el líquido. Pero la emergencia de lo mental en ciertos niveles de complejidad biológica es diferente. Según la posición emergente que está siendo considerada ahora, la conciencia es algo completamente nuevo.

A causa de que tal fenómeno de emergencia, incluso si es sistemático, permanece inexplicado, el ideal de inteligibilidad exige que consideremos con seriedad la alternativa de una respuesta reductiva al problema constitutivo, una respuesta que dé cuenta de la relación entre mente y cerebro desde el punto de vista de algo más básico en el orden natural. Si tal explicación fuera posible, explicaría la aparición de la vida mental en niveles complejos de organización biológica a través de un monismo general según el cual los constitutivos del universo tienen propiedades que explican no solo su carácter físico sino también el mental. Tom Sorell establece ese punto con claridad:

> Incluso si los mecanismos que produjeron la vida biológica, incluyendo la conciencia, son, en algún nivel, los *mismos* que los que operan en la evolución del universo físico, no se sigue que estos mecanismos sean físicos solo porque la evolución física precedió a la evolución biológica. Quizás se requiera algún concepto transfísico y transmental para aprehender ambos mecanismos. Esta conjetura presta atención a un ámbito llamado a veces por algo «monismo neutral» que se añade a las posiciones dualistas, materialistas e idealistas[32].

Sorell usa aquí «monismo neutral» para designar no solo una posición metafísica sino un tipo de teoría explicativa sistemática distinta del materialismo tradicional. Considerado solo metafísicamente, como respuesta al problema mente-cuerpo, el monismo sostiene que ciertos

estados físicos del sistema nervioso central son también necesariamente estados de conciencia, siendo su descripción física solamente una descripción parcial de ellas, desde fuera, por decirlo así. La conciencia no es en ese caso, como en la explicación emergente, un *efecto* de los procesos cerebrales que constituyen su condición física; más bien, estos procesos cerebrales son *por ellos mismos* más que físicos y lo incompleto de la descripción física del mundo se ejemplifica por lo incompleto de su descripción puramente física.

Pero ya que los organismos conscientes no están compuestos de un tipo especial de sustancia sino que pueden estar construidos, aparentemente, de cualquier materia del universo, convenientemente ordenada, se sigue que este monismo será universal. Todo, vivo o no, está constituido de elementos que tienen una naturaleza que es tanto física como no física, es decir, capaz de combinarse en todos mentales. Así esta explicación reductiva puede también ser descrita como una forma de panpsiquismo: todos los elementos del mundo físico son también mentales[33]. Sin embargo, el sentido en que son mentales se agota por la afirmación de que son tales para proporcionar una explicación reductiva de cómo sus combinaciones apropiadas constituyen organismos conscientes del tipo que conocemos. Cualesquiera consecuencias ulteriores de su carácter más-que-físico en el micro-nivel permanecen inespecíficas debido a lo abstracto de la propuesta.

5

Habiendo descrito la diferencia entre los dos tipos de respuesta al problema constitutivo, emergente y reductiva, permítaseme volver a la cuestión histórica, de nuevo con la conjetura de que el reduccionismo psicofísico es falso. La respuesta naturalista prevalente a la cuestión histórica es la versión materialista de la teoría evolucionista complementada con una explicación química especulativa del origen de la vida. La pregunta es: ¿qué alternativas hay a esta imagen si se rechaza el reduccionismo psicofísico?

La explicación histórica de cómo los organismos conscientes surgieron en el universo puede adquirir una de estas tres formas: será o causal

(apelando solamente a la causalidad eficiente gobernada por leyes), o teleológica, o intencional. (1) Una explicación histórica causal sostendrá que el origen de la vida y su evolución hasta el nivel de los organismos conscientes tiene su explicación última en las propiedades de los constitutivos elementales del universo, que son también los constitutivos de los organismos conscientes, junto con propiedades posteriores que pueden emerger como consecuencia de su combinación. (Si la explicación constitutiva de la conciencia no es emergente sino reductiva, entonces la explicación histórica causal será también completamente reductiva). (2) Una explicación teleológica sostendrá que, además de las leyes que gobiernan la conducta de los elementos en cada circunstancia, hay también principios de auto-organización o de desarrollo de la complejidad conforme avanza el tiempo que no son explicados por esas leyes elementales. (3) Una explicación intencional sostendrá que aunque el orden natural proporciona las condiciones constitutivas para la posibilidad de los organismos conscientes, como proporciona las condiciones de posibilidad de un avión a reacción, la realización de esa posibilidad se debió a la intervención de un ser (presumiblemente Dios) que unió los elementos constitutivos de la forma adecuada, quizás reuniendo el material genético que acabaría por último en la evolución de la vida consciente.

Ya que tanto una explicación constitutiva reductiva o emergente podría combinarse con cualquiera de los tres tipos de explicación histórica —causal, teleológica o intencional— hay seis opciones. Permítaseme decir algo sobre las explicaciones causales antes de volver a las otras posibilidades, que se alejan mucho más radicalmente de la forma usual de explicación científica.

Una explicación histórica causal puede combinarse con una explicación constitutiva emergente o con una reductiva. En la primera alternativa, la explicación histórica estará restringida a explicaciones puramente físicas del origen y evolución de la vida hasta el momento en el que los organismos alcanzaron la clase de complejidad que está asociada a la conciencia. Después de eso la historia sería tanto física como mental y, si el elemento mental emergente desempeñó una función causal independiente y no era simplemente epifenoménico, el proceso causal dejaría de ser estrictamente

reductivo. Pero tengo interés por la hipótesis de una historia causal físicamente reductiva que conduzca al menos a un punto en el cual acontezca la emergencia psicofísica, quizás en criaturas con sistema nervioso, quizás antes. Esta hipótesis preservaría la versión estándar de la evolución física sin alterar la marcha de la emergencia de la conciencia.

Antes discutí la cuestión de si una explicación física de la historia evolutiva unida a una teoría psicofísica no histórica podría explicar realmente la aparición de la conciencia y concluí que, a menos que hubiera algún vínculo ulterior entre la historia física y la teoría psicofísica, no aportaría un resultado inteligible incluso aunque fuera causalmente exacto. Presentaría la conciencia como un misterioso efecto colateral de la evolución biológica, quizás inevitable pero inexplicable en cuanto tal. Para explicar la conciencia, una historia evolutiva física tendría que mostrar por qué era probable que surgiesen aquellos organismos *del tipo que tienen conciencia*.

Eso sería posible si la teoría psicofísica que rige la emergencia de la conciencia revelase ser inseparable de la clase de organización física y funcionamiento de la vida animal cuyo desarrollo pretende explicar la historia evolutiva física mediante la selección natural[34]. Eso nos haría avanzar mucho a la hora de convertir la teoría evolucionista en una explicación de por qué existe la vida consciente. Implicaría que los organismos conscientes se han desarrollado mediante la selección natural precisamente en virtud de los tipos de características físicas que sistemáticamente dan lugar a la conciencia, según la teoría psicofísica de la emergencia. Esta es una opción seria. Tiene la desventaja de postular el hecho bruto de la emergencia, no explicable en términos de algo más básico, y por tanto esencialmente misterioso. Y descansa sobre la inmensa suposición de que una teoría física reductiva podría tener suficiente probabilidad para la aparición en el tiempo geológico del tipo adecuado de organismos físicos que desencadena esa emergencia. Pero pudiera considerarse en tanto que la explicación histórica requiere una pequeñísima alteración de la forma física prevalente de naturalismo mientras, sin embargo, reconoce la irreductibilidad de lo mental a lo físico.

Sin embargo, el otro tipo de explicación histórica causal, fundado en una teoría constitutiva reductiva más bien que en una emergente, explicaría en principio más. En un sentido diferente, pudiera incluso decirse que el punto de partida menos radical del reduccionismo materialista sería una concepción monista reductiva que sería tanto constitutiva como histórica; como teoría física aspira a serlo con respecto al mundo físico. La pregunta es si tiene sentido.

Una concepción comprehensivamente reductiva se ve favorecida por la creencia de que la propensión para el desarrollo de organismos con un punto de vista subjetivo debe haber estado allí desde el principio, tanto como la propensión a la formación de átomos, moléculas, galaxias y compuestos orgánicos debe haber estado allí desde el principio como consecuencia de las propiedades ya existentes de las partículas elementales. Si imaginamos una explicación que adquiere la forma de una versión ampliada del orden natural, con fenómenos locales complejos formados por composición desde elementos básicos universalmente disponibles, dependerá de alguna clase de monismo o pansiquismo más bien que de las leyes de la emergencia psicofísica que entran en el juego solamente de forma tardía.

Sin embargo, no está claro que esta clase de explicación reductiva pueda hacer inteligible esa conclusión en la forma en que la física de partículas o algo comparable vuelve ostensiblemente inteligible el carácter e historia cosmológica del mundo material no vivo. Las propiedades protopsíquicas de toda materia, en tal punto de vista, son postuladas solamente porque se necesitan para explicar la aparición de la conciencia en los niveles altos de complejidad orgánica. Además de que nada se conoce sobre ellas: son completamente indescriptibles y no tienen efectos locales predecibles, a diferencia de las propiedades físicas de los electrones y protones, que les permiten ser detectados individualmente. Así que no tenemos ni idea de cómo funcionaría tal explicación composicional. Sin algo inimaginablemente más sistemático, en la forma de una reducción, el pansiquismo no proporciona un nuevo y más básico punto de apoyo en la búsqueda de la inteligibilidad, un conjunto de principios básicos desde los cuales pueda verse que se sigan resultados más complejos. Ofrece

solamente la forma de una explicación sin ningún contenido y por tanto no parece ser un gran avance respecto a la alternativa emergente.

Pero la propuesta no es vana. En su forma esquemática, pre-socrática, esta clase de monismo intenta reconocer lo mental como una parte físicamente irreductible de la realidad mientras se separa de la forma básica de comprensión que ha resultado tan exitosa en la teoría física. Esto no es solo imitación intelectual; es fomentada por la conexión cercana entre mentes y cuerpos. Los organismos son complejos físicos cuya existencia y operación parece pedir una explicación reductiva y su existencia y operación parece amplia o completamente responsable de la existencia de la conciencia. Parece natural en consecuencia intentar incorporar la explicación de la conciencia a la misma estructura reductiva.

Por otra parte, la idea de reducir la mente a eventos o partículas mentales elementales no parece natural en el sentido en que sí lo es el atomismo físico. El marco espacio-temporal del mundo físico hace inmediatamente comprensible, geométricamente, la relación física parte-todo, pero no tenemos una idea clara comparable de una relación parte-todo de la realidad mental, ninguna idea de cómo los estados mentales en el nivel de los organismos pudieron resultar de las propiedades de los micro-elementos, y si estas propiedades son similares o diferentes en tipo a nuestros estados empíricos. Incluso un monismo mentalista tendría presumiblemente que encontrar las partes protomentales en un monismo —su contraparte— de las partes físicas del organismo y tendría que incluir una teoría de cómo se combinan en todos conscientes.

Resulta todavía más obscuro saber de qué manera las propiedades que explican cómo se constituyen los seres conscientes a partir de los elementos universales ayudarían a explicar cómo han surgido los seres conscientes, históricamente, en virtud de las leyes o principios que gobiernan la conducta de esos elementos. Si la teoría va a ser no solo constitutiva sino históricamente reductiva, entonces la naturaleza protomental de los elementos tiene que formar parte de la explicación de cómo empezó la vida y se desarrolló incluso antes de la aparición de los organismos animales.

Ya es parte natural de la concepción monista que los rasgos protomentales de los constitutivos básicos no son meramente pasivos sino

que son también necesariamente activos ya que se necesita explicar la inseparabilidad de lo activo y lo pasivo en la conciencia de los animales. Exactamente como la fenomenología y la conducta están internamente conectadas en la vida mental de los organismos, algo análogo debe ser verdad en el micro nivel si el monismo es correcto. Así, lo protomental tendrá implicaciones conductuales. Incluso más, si un monismo universal es correcto, significaría que estas conexiones psicofísicas son inquebrantables: uno no puede tener lo mental sin el aspecto físico o viceversa[35].

Pero esto no nos ayuda a imaginar una alternativa monista a la historia materialista del origen y evolución de la vida antes de la aparición de los organismos conscientes. Una vez que los organismos conscientes aparecen en escena, podemos ver cómo irá la cosa. Por ejemplo, un monismo reductivo implicaría que ciertas estructuras tienen necesariamente una experiencia visual en un sentido que inextricablemente combina fenomenología y capacidades de discriminación en el control de la acción y que no son posibles estructuras capaces del mismo control sin fenomenología. Si tales estructuras aparecieron en el menú evolutivo, presumiblemente reforzarían la eficacia de los organismos resultantes. De esa forma, lo protomental desempeñaría una función verdaderamente explicativa en la evolución biológica y no meramente epifenoménica.

Pero eso no explicaría por qué tales estructuras se formaron en primer lugar. Incluso si la posibilidad de un sistema visual está ya de alguna forma implicada por las propiedades de los elementos básicos, ¿cómo puede un monismo no materialista ayudar a explicar su aparición de hecho, a través del tiempo geológico? ¿Cómo podrían los mismos principios activos que explican la acción y la percepción en un organismo completamente formado dar razón también de la formación original de los organismos y la generación de mutaciones viables a través de la historia evolutiva? Esas preguntas son análogas a aquellas que se pueden plantear respecto de una teoría evolutiva reductiva materialista y le resultan exactamente tan difíciles de responder a una teoría no materialista.

Los mismos problemas habrá a la hora de explicar el origen de la vida y la disponibilidad de un suministro suficiente de mutaciones viables para que la selección natural funcione, es decir, suficiente para explicar la aparición

(ahora consciente) de la vida tal y como la conocemos. La clase de monismo o pansiquismo que se necesitaría para proporcionar una solución no emergente al problema constitutivo no hará más fáciles estas cuestiones históricas. Se asume que la química desempeña esta doble función en la explicación materialista estándar tanto de la operación del viviente como de la historia evolutiva de los organismos físicos, incluyendo el origen de la vida. Eso es algo altamente especulativo, pero un monismo hipotético que se ha ampliado para incluir la mente es mucho más especulativo ya que solamente dice que hay más sustancia básica del mundo de la que puede ser aprehendida por la física y la química. El objeto es organizar de nuevo la explicación de la evolución de los organismos animales para que dé cuenta no solo de su naturaleza física sino también de la naturaleza y funcionamiento de sus conciencias. Pero incluso si concluimos que las bases de la mente deben estar presentes en cada parte del universo, eso no ofrece ninguna pista de cómo las propiedades monistas que son la base de la conciencia en los organismos vivos conducen primero al origen de la vida y finalmente a la aparición de los sistemas conscientes dado el menú de mutaciones disponibles para la selección natural.

Nuestras creencias sobre las propiedades de los elementos físicos y sus constitutivos se basan en lo que se necesita para explicar sus conductas observables e interacciones y los resultados de su combinación en moléculas y estructuras más amplias. La forma materialista de naturalismo asume que la historia del universo desde el *big bang,* incluyendo el origen y la evolución de la vida, puede ser explicada por esas mismas propiedades. Esa es una conjetura muy grande, y una conjetura análoga tendría que ser la base de la hipótesis histórica del monismo reductivo si también se funda en las propiedades de los elementos necesarios para responder la cuestión constitutiva de una forma que incluya a la conciencia como un rasgo físicamente irreductible de ciertos organismos. ¿Por qué deberían estas propiedades hacer probable la aparición de tales organismos, partiendo de la materia inorgánica?

La idea de una respuesta reductiva tanto a la cuestión constitutiva como histórica resulta muy obscura. Busca una explicación más profunda y más unificada cósmicamente de la conciencia que la teoría emergente, pero al

coste de una mayor obscuridad y no ofrece ventaja evidente alguna respecto del problema histórico de la probabilidad.

6

Permítaseme comentar más brevemente las alternativas intencional y teleológica, cuyos atractivos se refuerzan por las dificultades a las que se enfrenta la explicación causal. También la respuesta a la cuestión constitutiva puede combinarse con una respuesta intencional a la cuestión histórica. Supongamos, por ejemplo, que la verdad constitutiva es reductiva. Entonces, si son posibles las explicaciones teístas, Dios podría haber llevado a cabo su propósito de crear seres conscientes o bien sacándolos de elementos con propiedades protopsíquicas o creando un universo con las condiciones iniciales altamente específicas apropiadas para hacer surgir a los seres conscientes a través de la evolución química y luego biológica, enteramente por leyes no teleológicas de interacción entre los elementos. La intención serviría en ese caso solamente como el marco externo de un sistema reductivo de causalidad eficiente. Para los teístas esto permanece como una opción.

Pero si estamos intentando imaginar una teoría laica, según la cual el desarrollo histórico de la vida consciente se explica completamente no por intervención sino como parte del orden natural, parece haber solamente dos alternativas: o este desarrollo depende enteramente de la causalidad eficiente, que opera en sus escenarios más tardíos a través de los mecanismos de la evolución biológica, o hay leyes naturales teleológicas que gobiernan el desarrollo de la organización a lo largo del tiempo, además de las leyes de tipo común que gobiernan el comportamiento de los elementos.

Esto es un retorno a la concepción aristotélica de la naturaleza, eliminada de la escena con el nacimiento de la ciencia moderna[36]. Pero estoy persuadido de que la idea de leyes teleológicas es coherente y completamente diferente de la idea de explicación por propósitos de un ser intencional que ajusta los medios a sus fines por elección. A pesar de la exclusión de la teleología de la ciencia contemporánea, ciertamente no se

debería descartar *a priori*. Formalmente la posibilidad de principios de cambio a lo largo del tiempo que tienden hacia ciertos tipos de resultados es coherente en un mundo en el que las leyes no teleológicas no son completamente deterministas[37]. Pero es esencial, si la teleología va a formar parte de un orden natural revisado, que sus leyes deban ser genuinamente universales y no solo la descripción de un solo proceso de búsqueda de fines. En tanto que conocemos un solo ejemplo de aparición y evolución de la vida, nos faltan bases para subsumirlo bajo leyes teleológicas universales, a menos que se puedan encontrar principios teleológicos operando con consistencia en niveles mucho más bajos. Tendría que haber tales leyes para que la teleología explicara auténticamente algo.

Cierto es que la idea de explicación teleológica está a menudo asociada a la idea ulterior de que los resultados tienen valor, de manera que sostener unos principios teleológicos particulares no es algo arbitrario. Lo que a su vez plantea la cuestión de si una explicación que apela al valor puede comprenderse independientemente de los propósitos de algún ser que lo anime. Una teleología sin propósito tendría que estar libre de valor o tendría que decir que el valor de ciertos resultados puede él mismo explicar por qué las leyes funcionan[38]. En ambos casos, la teleología natural significaría que el universo está racionalmente gobernado de más de una forma, no solamente a través de las leyes cuantitativas de la física que son la base de la causalidad eficiente, sino también a través de principios que implican que las cosas pasan porque están en un camino que lleva hacia ciertos resultados, notablemente, la existencia de organismos vivos y, por último, de los conscientes.

La opción teleológica es obscura en muchas formas. Tendré más que decir al respecto más tarde. La alternativa causal reductiva es igualmente obscura, pero si tuviera sentido, tendría el atractivo de una mayor unidad que la teleológica, porque significaría no solamente que los elementos en los que consiste el mundo natural tienen propiedades que acaban en organismos conscientes cuando se combinan convenientemente, sino que estas mismas propiedades no vuelven improbable que tales combinaciones se formaran de hecho por algún proceso gradual en el curso de la historia

cosmológica dado el tiempo disponible. Las cuestiones constitutiva e histórica se responderían entonces por referencia a un conjunto común de principios.

7

Hasta ahora he planteado el problema enfatizando la irreductibilidad de la experiencia consciente a lo físico. Pero he aludido al hecho de que la conciencia humana no es meramente pasiva sino que está penetrada, tanto en acción como en cognición, por la intencionalidad, la capacidad de la mente de representar el mundo y sus propios propósitos. Puede ser más controvertido declarar que la intencionalidad no puede realizarse en un universo puramente físico que decir que la conciencia no puede serlo. Sin embargo si, como creo, la intencionalidad, el pensamiento y la acción se resisten a la reducción psicofísica y pueden existir en las vidas de los seres que son también capaces de conciencia, entonces ellas también forman parte de lo que debe dar cuenta una explicación más amplia de lo mental. Este asunto lo iniciaremos en los siguientes capítulos. Creo que la función de la conciencia en la supervivencia de los organismos es inseparable de la intencionalidad: inseparable de la percepción, la creencia, el deseo y la acción y, finalmente, de la razón. La génesis de la estructura mental entera tendría que explicarse por principios básicos si se reconoce como parte del orden natural.

La filosofía no puede generar tales explicaciones; puede solamente señalar los lapsos que tienen y los obstáculos para construirlas con los materiales disponibles en el momento presente. Pero, a diferencia del dualismo clásico, sugiero que no deberíamos renunciar a la aspiración de encontrar una explicación naturalista integradora de un nuevo tipo. Tal teoría no puede ser abordada directamente. Requeriría muchos escenarios, a lo largo de un gran período de tiempo, comenzando con una cantidad enormemente mayor de información empírica sobre las regularidades en la relación entre estados conscientes y estados cerebrales en nosotros mismos y en los organismos cercanos. Solamente más tarde podría la hipótesis reductiva formularse sobre esta base de evidencia. Pero creo que tiene

sentido seguir la investigación no solo neurofisiológica sino también evolutiva, teniendo en mente una cierta meta utópica a largo plazo. Deberíamos buscar una forma de comprender que nos capacite para vernos a nosotros mismos y a otros organismos conscientes como expresiones específicas simultáneamente del carácter físico y mental del universo. Alguien podría objetar que la vida es ya bastante difícil de entender considerada puramente como fenómeno físico y que la mente puede esperar. Pero añadiendo el requisito de que cualquier teoría de la vida también tiene que explicar el desarrollo de la conciencia puede que el problema no empeore. Quizás, por el contrario, los rasgos añadidos del orden natural necesarios para explicar la mente contribuirán al final también a la explicación de la vida. Mientras más tiene que explicar una teoría, más poderosa tiene que ser.

CAPÍTULO 3

EL CONOCIMIENTO

1

Quiero ahora comenzar un tipo de argumento antirreduccionista diferente y ver sus consecuencias. La conciencia representa un problema para el reduccionismo evolucionista por su irreductible naturaleza subjetiva. Eso es verdad incluso en las formas más primitivas de conciencia sensorial tales como las que presuntamente se han encontrado en todos los animales. El problema que quiero tratar ahora se refiere a funciones mentales como pensamiento, razonamiento y valoración que están limitadas a los humanos, aunque incipientemente puedan encontrarse en otras pocas especies. Estas son las funciones que nos han capacitado para trascender la perspectiva del mundo de la vida inmediato que nos es dado por nuestros sentidos e instintos y para explorar la realidad objetiva más amplia de la naturaleza y del valor.

Daré por sentado que atribuir conocimiento a un ordenador es una metáfora y que solo un ser que tenga conciencia puede poseer las capacidades cognoscitivas de más alto nivel (dejando a un lado la cuestión de si su ejercicio puede ser inconsciente algunas veces). Eso implica que estas capacidades no pueden comprenderse solo mediante la ciencia física y que su existencia no puede ser explicada por una versión físicamente reductiva de la teoría evolucionista. Pero el problema que quiero plantear ahora va más allá de eso. Tiene que ver con la naturaleza de estas capacidades y cómo nos sitúan en el mundo. Lo que consideramos que estamos haciendo cuando pensamos sobre un asunto o sobre cómo

deberíamos actuar es algo que no se puede reconciliar con el naturalismo reductivo por razones distintas de las que implica la irreductibilidad de la conciencia. El problema no es solo la subjetividad del pensamiento, sino su capacidad para trascender la subjetividad y descubrir lo que es objetivo.

Pensar y razonar son correctos o incorrectos en virtud de algo independiente de las creencias del pensador e incluso independiente de la comunidad de pensadores a la que pertenece. Consideramos que tenemos la capacidad de formar creencias verdaderas sobre el mundo que nos rodea, sobre los ámbitos atemporales de la lógica y la matemática y sobre lo que es correcto hacer. No consideramos que estas capacidades sean infalibles, pero pensamos que se puede confiar habitualmente en ellas, en un sentido objetivo, y que pueden aportarnos conocimiento. La actitud natural interna de la vida humana cree que hay un mundo real, que muchas preguntas, ya sean fácticas o prácticas, tienen respuestas correctas y que hay reglas de pensamiento que, si las seguimos, nos conducirán a respuestas adecuadas a esas cuestiones. Se conjetura que seguir esas normas es responder correctamente a valores o razones que aprehendemos. La matemática, la ciencia y la ética se construyen sobre tales normas.

Es difícil dar sentido a todo esto desde el punto de vista del naturalismo tradicional. A menos que consideremos la mayor parte de esto como una ilusión, apunta a una ampliación de nuestra concepción del orden natural que incluya no solamente el origen de la conciencia fenomenológica —sensación, percepción y emoción— sino también el origen de nuestra capacidad activa para pensar por nosotros mismos más allá de estos puntos de partida. La pregunta es cómo entender la mente en su sentido integral como producto de la naturaleza o, más bien, cómo entender la naturaleza como un sistema capaz de generar mente.

El problema no surge en las formas básicas de la conciencia perceptiva, emocional y apetitiva que compartimos con muchos otros animales. Esas funciones mentales nos sitúan en una relación compleja con el mundo que nos rodea, pero parecen en principio susceptibles de una explicación evolucionista siempre y cuando sea de alguna manera transformada de la versión materialista en otra capaz de explicar la naturaleza consciente de esas funciones. Como indiqué en el último capítulo, si tales experiencias

pueden añadirse al menú evolucionista, sus funciones para capacitar a las criaturas para ir por el mundo, escapar de los peligros, encontrar alimento y cobijo y reproducirse, las harían potencialmente adaptativas y, por tanto, candidatas para la selección natural. La percepción y el deseo tienen que encontrar ciertos criterios de exactitud para hacer posible que las criaturas sobrevivan: tienen que capacitarnos para responder de forma similar a cosas similares y de forma diferente a cosas que son diferentes, huir de lo que es dañino y perseguir lo que es beneficioso. Para la mayoría de las criaturas la objetividad, sin embargo, no va más allá de eso. Sus vidas son vividas en un mundo de apariencias y la idea de una realidad más objetiva no tiene sentido.

Pero una vez reconocemos la distinción entre apariencia y realidad y la existencia de la verdad objetiva fáctica o práctica que va más allá de lo que la percepción, el apetito o la emoción nos dicen, la capacidad de criaturas como nosotros de llegar a tal verdad, o incluso de pensar sobre ella, requiere una explicación. Un aspecto importante de esta explicación será que hemos adquirido lenguaje y la posibilidad de comunicación interpersonal, justificación y crítica que el lenguaje hace posible. Pero la explicación de nuestra capacidad de adquirir y usar el lenguaje de estas formas presenta problemas del mismo orden porque el lenguaje es una de las facultades regidas normativamente más importantes. Adquirir un lenguaje es en parte adquirir un sistema de conceptos que nos capacita para entender la realidad.

En este punto voy a obviar todos los problemas mencionados antes sobre la probabilidad del origen de la vida y la suficiencia de la mutación aleatoria y la selección natural para dar razón de la historia evolutiva fáctica de la vida en la Tierra. El problema que quiero suscitar se mantiene incluso si esos otros se pueden resolver por la evolución de las plantas y de los animales más rudimentarios. También supondré que la teoría evolucionista puede ser reformulada de forma que sea consistente con el antirreduccionismo, para que pueda explicar la aparición de la conciencia. La cuestión que ahora quiero plantear es si nuestras capacidades cognoscitivas pueden caber en el marco de una teoría evolutiva que no sea exclusivamente materialista pero que retenga la estructura darwiniana. Es

una cuestión hipotética ya que no puede haber tal teoría. Pero hablaré como si la hubiera.

El problema tiene dos aspectos. El primero afecta a la probabilidad de que el proceso de selección natural pueda haber generado criaturas con la capacidad de descubrir con la razón la verdad sobre una realidad que se extiende vastamente más allá de las apariencias iniciales, como consideramos que hemos hecho y continuaremos haciendo colectivamente en ciencia, lógica y ética. ¿Se puede creer que la selección por eficacia adaptativa en el pasado prehistórico ha fijado las capacidades que son efectivas en las búsquedas teóricas que fueron inimaginables en ese tiempo? El segundo problema es la dificultad de comprender naturalistamente la facultad de la razón que es la esencia de esas actividades. Comenzaré considerando una posible respuesta al primer problema antes de pasar al segundo, que es particularmente intratable.

2

El primer problema surge solamente si se presupone el realismo. Queremos saber, por ejemplo, la probabilidad de que la evolución haya dado a algunos seres humanos la capacidad de descubrir, y a otros seres humanos la capacidad de comprender, las leyes de la física y la química. Si no hay un mundo físico real independiente del juicio, ni verdades matemáticas independientes del juicio, ni verdades de la ética y la razón práctica independientes del juicio, entonces no existe el problema de explicar cómo somos capaces de aprenderlas. En un punto de vista antirrealista, la verdad, científica o moral, depende de nuestras respuestas sistemáticas cognoscitivas o conativas más bien que ser algo independiente a lo que nuestras respuestas puedan o no puedan conformarse. Los «mundos» en cuestión son, todos, solo construcciones humanas. En ese caso una explicación de cómo se formaron esas respuestas —incluyendo nuestras teorías científicas— no tendrá que explicar su corrección objetiva para ser aceptable (aunque tendrá que explicar su coherencia interna).

Un antirrealismo de este tipo es una opción más seria para el caso de la moral que para el de la ciencia. Uno puede comprensiblemente sostener que

el realismo moral es implausible porque la teoría evolucionista es la mejor explicación actual de nuestras facultades y una explicación evolucionista no puede depender de si seremos capaces de descubrir verdades morales independientes del juicio, si es que hay tal cosa[39]. Pero sería inadecuado abandonar el realismo científico por motivos análogos, porque uno tendría entonces que ser también antirrealista respecto de la teoría de la evolución. Esto significaría que la teoría evolucionista es inconsistente con el realismo científico y no puede ser comprendida realistamente, lo que parece una conclusión excesivamente fuerte. Hay algo raro hasta el punto de la incoherencia en adoptar el naturalismo científico como fundamento del antirrealismo en la ciencia natural.

Si abandonamos la creencia en el realismo, la mayor esperanza para la validez de una respuesta naturalista al primer problema sería que la teoría evolucionista, y en particular la psicología evolucionista, fuera capaz de dar razón del éxito de nuestras capacidades cognoscitivas. Esta es la aspiración de la epistemología naturalista para el conocimiento fáctico. Su objetivo sería explicar cómo las capacidades mentales innatas que fueron seleccionadas por su valor adaptativo inmediato son también capaces de generar, a través de una historia evolutiva cultural ampliada, teorías verdaderas sobre un orden natural gobernado por leyes que no fuera adaptativamente necesario comprender antes. La explicación evolucionista tendría que ser indirecta ya que el conocimiento científico no tiene ninguna función en la selección de las capacidades que lo generan.

Ese así-fue-la-historia sería a grandes rasgos como sigue. En la vida salvaje no hay percepción ni condicionamiento operante que no tenga valor de supervivencia. La capacidad de generalizar de la experiencia y permitir que estas generalizaciones, o expectativas generales, sean confirmadas o no por la experiencia es también adaptativa. Así, hay una disposición básica a mantener la consistencia lógica de la creencia modificándola cuando surgen inconsistencias. Un paso siguiente, muy importante, sería la capacidad de corregir las apariencias individuales, no solo por referencia a otras apariencias propias que entren en conflicto, sino también por referencia a cómo aparecen las cosas a otros perceptores. Eso requiere el reconocimiento de otras mentes, una capacidad de un potencial adaptativo

obvio. El alcance de estas capacidades puede ser bastante ampliado y deliberadamente ejercido con la ayuda del lenguaje, que también permite que el conocimiento sea creado colectivamente, acumulado y transmitido. Con el lenguaje podemos contemplar nuestras mentes y compartir con los demás posibles alternativas y decidir entre ellas sobre el fundamento de su consistencia o inconsistencia con posteriores observaciones. Las teorías científicas complejas que comportan predicciones empíricas son, por tanto, extensiones de la capacidad altamente adaptativa de aprender de la experiencia, nuestra y de los otros.

Esta historia depende de la suposición de un origen biológico de la capacidad de representación no perceptual del lenguaje, que culmina en la capacidad de captar estructuras abstractas lógicamente complejas. No es fácil decir cómo decidir si esto podría ser una manifestación de capacidades que tienen valor de supervivencia en la vida diaria prehistórica. En vistas de la sofisticación matemática de las modernas teorías físicas parece altamente improbable, pero quizás se podría defender esa tesis.

Es posible contar un así-fue-la-historia paralelo sobre la compatibilidad entre teoría evolucionista y realismo moral. No estoy pensando en la común apelación a la sociobiología con su esencial interpretación nepotista de las disposiciones altruistas innatas. No estoy pensando tampoco en la explicación a través de la selección grupal de disposiciones para la cooperación en las criaturas sociales[40]. Más bien tengo en mente el descubrimiento de principios generales de valor por medios racionales análogos a los usados habitualmente. Comenzando desde una comprensión de los deseos y aversiones innatas como impresiones inmediatas de valor —de lo que es bueno o malo para nosotros mismos o nuestra parentela— habría que descubrir un ámbito normativo más amplio gobernado por principios, o ámbito de la razón práctica, en el que esos valores aparentes puedan evolucionar a través de la capacidad de generalizar y la tendencia a evitar la inconsistencia.

La generalización conduciría al reconocimiento del valor en posibles experiencias futuras, en los medios para ellas, y en las vidas de otras criaturas además de nosotros mismos. Esos valores no son propiedades extra de bondad o maldad, sino solo verdades tales como la siguiente: si

algo que hago causa sufrimiento a otra criatura, eso pesa para no hacerlo. Puedo llegar a ver que esto es verdad generalizando desde el evidente disvalor de mi propio sufrimiento y, una vez que reconozca esa verdad más general, mis motivos cambiarán. Si hay normas de conducta generales objetivas, esta clase de pensamiento nos permitiría descubrirlas incluso si no son más innatas que las leyes de la física. Como con la ciencia, el proceso de descubrimiento sería imposible sin el lenguaje, la comunicación interpersonal y la memoria cultural. En ambos casos, aunque las capacidades básicas empleadas son adaptativas en su forma simple, nos permitirían trascender nuestros puntos de partida para descubrir amplios ámbitos de verdad con independencia de si tal conocimiento refuerza la adaptación.

Todo esto es altamente improbable, pero no más que muchas otras especulaciones evolucionistas. Requiere que las mutaciones y cualesquiera otras cosas que puedan ser fuentes de la variación genotípica generaran no solo estructuras físicas sino, fenomenológicamente, deseo y aversión, conciencia de otras mentes, representaciones simbólicas y consistencia lógica, desempeñando todo ello funciones esenciales en la producción de conducta. Siempre y cuando podamos conjeturar alguna solución global al problema mente-cuerpo que lleve a todo esto, el resto de la historia sugiere que el conocimiento de la verdad científica objetiva y de la verdad moral —debería haber tales cosas— podría ser resultado del ejercicio de capacidades que, en aplicaciones más mundanas, no son enemigas de la supervivencia. Que sea improbable podría no resultar un problema insuperable siempre y cuando aceptemos el marco evolucionista mismo como probable.

3

Sin embargo, incluso si suponemos que es verdadera alguna explicación evolucionista de este tipo, hay un problema ulterior que dificulta que podamos pensar de esta manera sobre las capacidades básicas del razonamiento. Surge si contrastamos la actitud que razonablemente adoptamos hacia nuestro sistema perceptivo y apetitivo con la actitud que

tomamos hacia nuestro razonamiento. Esto llevará al segundo problema indicado arriba: la dificultad de comprender la razón naturalistamente.

Si suponemos que hay alguna forma de incluir la conciencia en la narración evolucionista, entonces podemos entender nuestro sistema visual como el sistema visual de otras especies, ya que ha sido formado por selección natural. Lo específico de la visión humana responde a aspectos del mundo que han sido importantes en las vidas de nuestros antepasados. Ello nos permite continuar confiando en la evidencia inmediata de nuestros sentidos aunque reconozcamos que algunas veces será engañosa, selectiva o distorsionada y que porta los caracteres de la particular biología de nuestros ancestros.

Algo similar es posible decir de nuestras actitudes hacia los juicios intuitivos de probabilidad, o hacia algunos de nuestros juicios intuitivos de valor (el deseo de venganza, por ejemplo). Podemos llegar a entender esas intuiciones como burdas pero útiles respuestas irreflexivas formadas por la selección natural para conseguir reforzar la eficacia en las circunstancias en las cuales vivieron y murieron nuestros ascendientes. Simultáneamente podemos reconocer que pueden necesitar corrección o inhibición. La autoconciencia evolutiva es un rasgo común de nuestras actitudes reflexivas hacia nuestras naturales inclinaciones de hambre, miedo, placer, ira, etc.

Pero cada vez que adoptamos tal actitud objetiva y razonable hacia nuestras disposiciones innatas, estamos implícitamente pensando de una forma a la cual no aplicamos esa actitud objetiva. Cuando confiamos en sistemas de medida para corregir la percepción o en cálculos de probabilidad para corregir las expectativas intuitivas, o un razonamiento moral o prudencial para corregir los impulsos instintivos, consideramos estar respondiendo a razones sistemáticas que justifican por ellas mismas nuestras conclusiones y que no obtienen su autoridad de sus orígenes biológicos[41]. Ellas *no podrían* ser confirmadas de esa manera. No obtienen su autoridad de sus orígenes culturales tampoco; por el contrario, la historia cultural que se ha desarrollado se valida como ejemplo de progreso solo por el hecho de que ha conducido a esos métodos de incrementar la exactitud de nuestros juicios.

Confiar en la propia visión y confiar en la razón propia se parecen en un aspecto: en ambos casos la confianza es inmediata. Cuando veo un árbol no infiero su existencia de mi experiencia más que cuando infiero la corrección de una inferencia lógica del hecho de que no puedo evitar creer en la conclusión. Sin embargo hay una diferencia crucial: en el caso perceptivo puedo reconocer que puedo estar equivocado, pero por reflexión, incluso si me pienso a mí mismo como el producto de la selección natural darwiniana, estoy a pesar de ello justificado para creer la mayor parte de las veces la evidencia de mis sentidos porque es consistente con la hipótesis de que una representación exacta del mundo que me rodea nace de sentidos formados por evolución para servir a esa función. Esa no es una refutación del escepticismo radical ya que la teoría evolucionista, como toda la ciencia, depende de la evidencia de los sentidos. Pero proporciona una imagen coherente de mi lugar en el mundo que es consistente con la confianza habitual en tal evidencia.

Por el contrario, en el caso del razonamiento, si es suficientemente básico, lo único que tengo que pensar es que he captado la verdad directamente. No puedo ir atrás de una inferencia lógica y reconfirmarla con la reflexión de que la confianza en los procesos de mi pensamiento lógico es consistente con la hipótesis de que la evolución los ha seleccionado por su exactitud. Eso debilitaría drásticamente la reivindicación lógica. Es más, en la formulación de esa explicación, como en la explicación paralela de la confianza en los sentidos, los juicios lógicos de consistencia e inconsistencia tienen que darse sin esas reservas como aprehensiones directas de la verdad. No es posible pensar: «La confianza en mi razón, incluyendo mi confianza en *este mismo juicio,* es razonable porque es consistente por tener una explicación evolutiva». Por tanto cualquier explicación evolutiva del lugar de la razón presupone la validez de la razón y no puede confirmarse sin circularidad.

Al final, el intento de entenderse a uno mismo en términos evolucionistas y naturalistas debe tocar fondo en algo que es afirmado como válido en sí mismo, algo sin lo que la comprensión evolucionista no sería posible. El pensamiento nos traslada más allá de la apariencia a algo que no podemos considerar simplemente como una disposición fundada

biológicamente y cuya fiabilidad podemos determinar sobre otros fundamentos. No es suficiente pensar que *si* hay verdades lógicas, la selección natural pudiera muy bien haberme dado la capacidad de reconocerlas. Ese no puede ser mi fundamento para confiar en mi razón porque incluso ese pensamiento confía en la razón en una forma anterior.

Podemos suponer que las capacidades que nos habilitan para ir mucho más allá de nuestras disposiciones innatas de representar y responder al mundo han aparecido en un ancestro y entonces han sido preservadas en las siguientes generaciones. La aparición de estas capacidades tiene que poder integrarse en el proceso evolutivo en que, al menos, no son perjudiciales para la eficacia adaptativa, de manera que no se extingan por selección natural. Eso parece muy plausible. Pero si estoy en lo cierto en pensar que no podemos considerarlas meramente como disposiciones instintivas ulteriores, se necesita alguna otra explicación de lo que son esas capacidades.

Exactamente igual como la conciencia no puede explicarse como una simple ampliación o complicación de la evolución física, de igual forma la razón no puede explicarse como una simple ampliación o complicación de la conciencia. Explicar nuestra racionalidad requerirá, además de lo que se necesita para explicar nuestra conciencia y sus formas evidentemente adaptativas, algo de un nivel diferente. La razón puede llevarnos más allá de las apariencias porque tiene completa validez general más que utilidad meramente local. Si la tenemos, reconocemos que no puede ser ni confirmada ni debilitada por una teoría de sus orígenes evolutivos ni por ningún otro punto de vista externo a ella misma. No podemos distanciarnos de ella. Esa fue la perspectiva de Descartes.

Si la razón apareció en el menú evolutivo, podría haber probado su valor adaptativo localmente. Luego, con la ayuda del despliegue y desarrollo cultural podría haberse elevado a su posición actual de autoridad crítica, corrigiendo y frecuentemente desautorizando a los más viejos impulsos de la percepción, el instinto y la intuición y no estando sujeta a corrección por nadie más. Su afianzamiento y final soberanía sobre instintos más viejos es comprensible, pero solamente si podemos comprender cómo tal cosa puede existir en absoluto.

Este es el segundo problema: ¿Cuál es la facultad que nos capacita para escapar del mundo de la apariencia presentado por nuestras disposiciones innatas prerreflexivas al mundo de la realidad objetiva? Y, ¿qué, además de la conciencia, tenemos que añadir a la historia biológica para dar sentido a tal facultad?

Lo distintivo de la razón es que nos conecta con la verdad directamente. La percepción nos conecta con la verdad solo indirectamente. Cuando veo un árbol lo veo porque está allí, pero no solo porque está allí. La percepción no es una forma de conocimiento: no capto la presencia del árbol inmediatamente, aunque pueda parecer anterior a la reflexión. Más bien soy consciente de él porque el árbol causa un efecto mental en mí en virtud de la naturaleza de mi sistema visual, que podemos suponer que ha sido formado por la selección natural para reaccionar de esa forma a la luz reflejada por los objetos físicos. Tener tal sistema, junto con otras disposiciones perceptivas y motivacionales, me capacita para sobrevivir en el mundo. Así es, solamente en un sentido complicado e indirecto, que cuando veo un árbol lo veo porque está allí.

Pero supongamos que observo una contradicción entre mis creencias y «veo» que debo abandonar al menos una de ellas. (Estoy conduciendo hacia el sur por la mañana temprano y el sol sale por mi derecha.) En ese caso, veo que las creencias contradictorias no pueden ser todas verdaderas y lo veo simplemente porque es así. Lo capto directamente. No es adecuado decir que, enfrentado a una contradicción siento la urgente necesidad de alterar mis creencias para escapar de ella, lo que se explica por el hecho de que evitar las contradicciones, como evitar serpientes y precipicios, afianzaba la eficacia adaptativa de mis ancestros. Esa sería una explicación indirecta de cómo la imposibilidad de la contradicción explica mi creencia de que no puede ser verdadera. Pero incluso si alguno de nuestros ancestros fuese presa de fobias e instintos meramente lógicos, nosotros hemos ido más allá de eso: rechazamos una contradicción solo porque vemos que es imposible y aceptamos una consecuencia lógica solo porque vemos que es necesariamente verdadera.

En la percepción ordinaria, somos como mecanismos gobernados por un (rudo) algoritmo que preserva la verdad. Pero cuando razonamos somos como un mecanismo que puede ver que el algoritmo que sigue preserva la verdad. Algo ha pasado que ha puesto nuestras mentes en contacto inmediato con el orden racional del mundo o, al menos, con los elementos básicos de ese orden y que puede usarse para lograr mucho más. Eso nos capacita para tener conceptos que muestren la compatibilidad o incompatibilidad de creencias particulares con hipótesis generales. Tenemos que comenzar considerando nuestras impresiones prerreflexivas como un punto de vista del mundo parcial y como en una perspectiva, pero estamos capacitados para usar la razón y la imaginación con el propósito de obtener candidatos de una concepción más amplia que pueda contener y dar cuenta de esa parte. Esto rige en el ámbito del valor tanto como en el del hecho. El proceso es altamente falible, pero no podría ni siquiera intentarse sin este núcleo de auto-evidencia del cual depende todo menos el razonamiento cierto. En la crítica y corrección del razonamiento, la corte final de apelación es siempre la razón misma.

Lo que esto significa es que si esperamos incluir la mente humana en el orden natural, tenemos que explicar no solamente cómo influye la conciencia en la percepción, la emoción, el deseo y la aversión, sino también el control consciente de la creencia y la conducta en respuesta a la conciencia de las razones —la evitación de la inconsistencia, la subsunción de casos particulares bajo principios generales, la confirmación o no de principios generales por observaciones particulares, etc. Esto es lo que le permite a uno guiarse por la verdad objetiva y no por las impresiones propias. Es un tipo de libertad, la libertad que la conciencia reflexiva nos da del gobierno de las disposiciones perceptivas y motivacionales innatas y del condicionamiento.

Las criaturas racionales pueden alejarse de estas influencias e intentar tomar sus propias decisiones. Dejo aparte la cuestión de si esta clase de libertad es compatible o incompatible con el determinismo causal, pero en verdad parece que es algo que no se puede someter a un análisis puramente físico y por tanto, como a las formas más pasivas de conciencia, no se le puede dar tampoco una explicación puramente física.

Si decido, cuando el sol sale por mi derecha, que debo estar conduciendo hacia el norte en lugar de hacia el sur, es porque reconozco que la creencia de que estoy conduciendo hacia el sur es inconsistente con esa observación, además de lo que sé acerca de la dirección de rotación de la tierra. Abandono la creencia porque reconozco que no podría ser verdadera. Si deposito dinero en un plan de pensiones porque el ingreso futuro que genera me será más valioso entonces de lo que pueda gastar ahora, actúo porque veo que es bueno hacerlo. Si me opongo a la abolición del impuesto de sucesiones es porque reconozco que el diseño de los derechos de propiedad debería ser sensible no solo con la autonomía sino también con la equidad. Como dice el dicho, obro en el espacio de las razones.

La aparición de la razón y del lenguaje en el curso de la historia biológica parece, desde el punto de vista de las formas disponibles de explicación, algo radicalmente emergente, si, como supongo, no puede ser comprendido conductistamente. Como la conciencia, presenta problemas tanto de explicación constitutiva como histórica. Apareció mucho después de la emergencia de las criaturas conscientes, incluso parece también ser esencialmente un desarrollo de la conciencia y debería ser comprensible como parte de esa historia. Como la conciencia, la razón es inseparable de la vida física de los organismos que la tienen, ya que actúa sobre el material proporcionado por la percepción y el deseo natural y controla la acción tanto directa como indirectamente. Cualquier comprensión de ella transformará nuestra comprensión de los organismos físicos y también de su desarrollo.

El gran giro cognoscitivo es una ampliación de la conciencia desde la forma de perspectiva que tiene en las vidas de las criaturas particulares a una forma objetiva, abarcadora del mundo, que existe tanto individual como intersubjetivamente. Fue originalmente un proceso biológico evolutivo y en nuestra especie ha llegado a ser también un proceso cultural colectivo. Cada una de nuestras vidas es una parte del largo proceso de un universo despertándose gradualmente y llegando a ser consciente de sí mismo.

Esto es, entonces, lo que una teoría del todo tiene que explicar: no solo la emergencia desde un universo carente de vida de organismos que se reproducen y su desarrollo por evolución a una cada vez mayor complejidad funcional; no solo la conciencia de alguno de esos organismos y su función central en sus vidas; sino también el desarrollo de la conciencia en un instrumento de transcendencia que puede captar la realidad objetiva y el valor objetivo.

Se pueden suponer ciertas cosas si existe la razón. Primero, hay verdades objetivas de diferentes clases, independientes de la mente: las verdades fácticas sobre el mundo natural, incluyendo las leyes de la ciencia; las verdades eternas y necesarias de la lógica y las matemáticas; y verdades valorativas y morales. Segundo, por comenzar de la forma en que las cosas nos aparecen inicialmente, podemos usar la razón colectivamente para alcanzar creencias justificadas sobre alguna de esas verdades objetivas, aunque alguna de esas creencias probablemente estará equivocada. Tercero, estas creencias, combinadas, pueden directamente influir en lo que hacemos. Cuarto, estos procesos de descubrimiento y motivación, en tanto que mentales, son inseparables de los procesos físicos del organismo.

Es una verdad trivial que si hay organismos capaces de razón, la posibilidad de tales organismos debe haber existido desde el principio. Pero si creemos en un orden natural, entonces algo del mundo que dio lugar a seres racionales debe explicar tal posibilidad. Es más, para explicar no simplemente la posibilidad sino el hecho de los seres racionales, el mundo debe tener propiedades que no conviertan su aparición en un completo accidente: de alguna forma la posibilidad debe haber estado latente en la naturaleza de las cosas. Así, tenemos la necesidad tanto de una explicación constitutiva de en qué pudiera consistir la racionalidad como de una explicación histórica de cómo surgió; y ambas explicaciones deben ser consistentes con nuestro ser, entre otras cosas, organismos físicos. La comprensión de los organismos biológicos y sus historias evolutivas tendría que ampliarse para incorporar esta carga explicativa adicional; como he

argumentado, debe ampliarse más allá del materialismo para incorporar la explicación de la conciencia.

Tal explicación completaría la búsqueda de la inteligibilidad mostrando cómo el orden natural tiende a generar seres capaces de comprenderlo. Pero los obstáculos parecen enormes. A la luz del carácter formidable de la razón, es difícil imaginar cómo podría ser viable una explicación naturalista de ella, ya sea constitutiva o histórica.

En el capítulo anterior, exploré la posibilidad de una explicación reductiva de la conciencia fundada en alguna forma de monismo universal o pansiquismo. Esto se ejemplificaba en el reduccionismo físico alentado por la biología molecular, pero con un fundamento metafísico ampliado en el que lo físico y lo mental eran ontológicamente inseparables. Aunque sería una desviación radical desde el punto de vista de la naturaleza del materialismo reinante, el monismo requerido para una explicación reductiva pero no físicamente reduccionista de la conciencia parece concebible. En respuesta a la cuestión constitutiva, la idea de que un sujeto de conciencia complejo pudiera construirse de elementos protomentales mínimos que son unificados simultáneamente de alguna forma en un organismo y un yo, tiene bastante potencial para merecer consideración. Visto como alternativa a una igualmente especulativa emergencia de la conciencia en los niveles altos de la organización física, parece relativamente creíble a pesar de los importantes problemas de la relación mental parte-todo.

Sin embargo, una explicación reductiva de la razón, netamente desde el punto de vista de las propiedades de los constitutivos elementales de los que están hechos los organismos, es incluso más difícil de imaginar que una explicación reductiva de la conciencia. La racionalidad, más incluso que la conciencia, parece necesariamente un rasgo del funcionamiento del sujeto consciente íntegro y no puede concebirse, ni incluso especulativamente, como compuesto de incontables átomos de racionalidad en miniatura. La metáfora de la mente como un ordenador formado de un enorme número de transistores que semejan homúnculos no servirá para el propósito, porque omite la comprensión del contenido y fundamentos del pensamiento y la acción esenciales a la razón. Podría explicar el *output* conductista, pero no la comprensión. Por estas razones, una respuesta holística o emergente a la

cuestión constitutiva es más probable que una reductiva cuando vamos de los organismos físicos a la conciencia y a la razón. Esto querría decir que la razón es una facultad irreductible, del tipo de mente consciente completamente formada que existe en los animales superiores, y que no puede ser analizada en la actividad de las partes protomentales de la mente, en la forma en que quizás la sensación pueda serlo.

Pero permanece la cuestión histórica. Incluso si ocurre algo enteramente nuevo cuando el cerebro consciente alcanza un determinado tamaño y nivel de complejidad, una explicación de la existencia de tal complejidad será adecuada solamente si también explica la existencia de la razón como tal. (Esto es paralelo a las exigencias de una explicación de la conciencia como tal expuestas en el último capítulo.) Supongamos que tenemos razón porque nuestros cerebros han alcanzado el nivel de complejidad en el que la razón emerge. Si esta va a ser una explicación que justifique que la aparición de la razón no es un completo accidente, debe de alguna forma explicar no solo la complejidad física misma sino la aparición del mismo tipo de complejidad que es condición de emergencia de la razón. Esto no sería necesario si uno considera a la razón como algo que ocurre por chiripa, un puro efecto colateral de otros desarrollos cerebrales. Pero si eso no es aceptable, entonces una explicación de la razón tendría que explicar la probabilidad de la aparición de sus condiciones biológicas *qua* condiciones de la razón, es decir, bajo esa descripción. Las diversas posibilidades en este asunto están demasiado abstractamente descritas para permitir cualquier especulación sobre si una explicación causal reductiva pudiera hacer esto, pero si la emergencia es la respuesta correcta a la cuestión constitutiva de la razón, puede ser que la cuestión histórica requiera o una solución teleológica o una intencional.

6

He abierto la posibilidad de principios teleológicos como parte del orden natural en el capítulo previo. La explicación teleológica puede tener graves problemas, pero en este caso no son más graves que los de sus alternativas, de manera que no habría que dejar de considerarla posible. La

evolución de la mente es parte de un solo y largo proceso de ascendencia evolutiva. Es el último escenario de la evolución de los organismos físicos, algunos de los cuales se rigen ahora ampliamente por el pensamiento. Si somos escépticos sobre una explicación intencional (teísta) de la existencia de la razón, y no podemos encontrar sentido a una reduccionista causal, es natural especular que algunas tendencias hayan estado trabajando durante todo el tiempo en esta dirección. Si la física sola o incluso un monismo no materialista no pueden dar razón de los últimos escenarios de nuestra historia evolutiva, no deberíamos suponer que pueden darlo de los escenarios más tempranos. En verdad, cuando retrocedemos suficientemente lejos, al origen de la vida —de sistemas auto-replicantes capaces de apoyar la evolución por selección natural—, los que se ocupan de investigar el tema reconocen que están muy lejos de formular una hipótesis explicativa viable del tipo materialista tradicional. Aunque suponen que debe haber una explicación tal, puesto que la vida no puede haber surgido puramente por casualidad[42].

Esa suposición puede estar fundada en una confusión. En un artículo importante, Roger White ha argumentado que la búsqueda de una explicación del origen de la vida desde el punto de vista de los principios no intencionales de la física y la química —una explicación que revelara que el origen de la vida no es meramente casualidad sino algo que era de esperar o, al menos, no sorprendente que ocurriera— está probablemente motivada por el sentido de que la vida no puede ser una cuestión de azar, porque suena mucho a como si fuera el producto de un diseño intencional. Pero la hipótesis del diseño intencional se descarta como no científica. Así, resulta natural concluir que la única forma que le queda a la vida para no ser una casualidad es estar constituida por medio de la ley física. Pero como White señala, esta inferencia es ilegítima. Esto es lo que dice:

> La cadena de razonamiento es similar a lo siguiente. Que los sistemas moleculares replicantes parecen ser diseñados por un agente es suficiente para convencernos de que no surgieron por casualidad. Pero en el razonamiento científico deben preferirse las explicaciones no intencionales, si es posible (alguno diría a toda

> costa), a las intencionales, de ahí la motivación para encontrar una explicación no intencional de la vida.
>
> Debería quedar claro, sin embargo, que incluso concediendo que es apropiada una preferencia por explicaciones no intencionales, esta línea de razonamiento está confundida. En general, si B_1 [la hipótesis de que el proceso que condujo a *S* estaba intencionalmente sesgado] eleva la posibilidad de *S*, entonces *S* confirma a B_1 al menos hasta algún grado, y puede así desconfirmar *C* [la hipótesis de la casualidad]. Pero no se sigue que *S* confirme B_N ni una tilde [la hipótesis de que el proceso no estaba intencionalmente sesgado]. *S* confirma B_N *solamente si* B_N eleva la posibilidad de *S*. Si la razón por la que dudamos de la Hipótesis de la Casualidad es que sospechamos que la vida se debe en parte a un agente inteligente, esto por sí mismo no nos da razón para esperar que haya una explicación no intencional de la vida. Si por reflexión no encontramos aceptable la hipótesis del sesgo intencional, entonces nos quedamos enteramente sin motivo para dudar de que la vida surgió por casualidad[43].

Gran parte del artículo de White se ocupa de argumentar que la aparición de la vida no pueda esperarse más bajo la asunción de B_N —la hipótesis del sesgo no intencional— que bajo la conjetura de la casualidad. Eso es porque si hay un sesgo no intencional hacia cierto resultado derivado de una ley física sin propósito, podría ser un sesgo hacia cualquier tipo de resultado, así que no puede hacer más probable la aparición de la vida que cualquier otra cosa. Como dice White:

> Lo que hace que ciertas configuraciones moleculares se destaquen de entre la multitud de posibilidades parece ser que ellas son capaces de desarrollarse en algo que nos resulta más bien maravilloso, a saber, un mundo de criaturas vivas. Pero no hay razón concebible para que fuerzas ciegas de la naturaleza o atributos

físicos de la misma condición deban estar sesgadas hacia lo maravilloso[44].

Por el contrario, el sesgo intencional está restringido como hipótesis por algunas conjeturas burdas sobre los motivos que dan lugar a las intenciones. (Entonces uno no puede reclamar ningún resultado *S*, aunque sea aleatorio o arbitrario, que evidencie que el proceso que condujo a él estaba intencionalmente sesgado, simplemente por el motivo de que se vuelve probable por la hipótesis de que fue producido por un ser que quiso precisamente que ocurriera *S*. Cualquier argumento del diseño depende de suposiciones generales más restrictivas sobre qué clase de cosas pudiera querer producir un diseñador.)

Me siento atraído por una cuarta alternativa, la teleología natural o sesgo teleológico, como una explicación de la existencia de las posibilidades biológicas sobre las que la selección natural puede obrar. Creo que la teleología es una alternativa naturalista que es distinta de todas las otras tres explicaciones candidatas: casualidad, creacionismo o ley física no dirigida. Para evitar el error que White encuentra en la hipótesis del sesgo no intencional, la teleología tendría que ser restrictiva en lo que la hace probable, sin hacerla depender de intenciones o motivos. Esto tendría que involucrar probablemente alguna concepción de un incremento de valor mediante las posibilidades ampliadas proporcionadas por las formas más altas de organización hacia las que tiende la naturaleza: cualquier resultado no podría calificarse como un *telos*. Eso daría valor a un fin explicativo, pero no a uno que se realizase por propósitos o intenciones de un agente. La teleología significa que además de la ley física de tipo común hay otras leyes de la naturaleza que están «sesgadas hacia lo maravilloso».

La idea de teleología como parte del orden natural va en contra de la forma acreditada de explicación que ha caracterizado a la ciencia desde la revolución del siglo XVII. La teleología significaría que algunas leyes naturales, a diferencia de todas las leyes científicas básicas descubiertas hasta ahora, son temporalmente históricas en su operación. Las leyes de la física son todas ecuaciones que especifican relaciones universales que funcionan en todo tiempo y lugar entre cantidades matemáticamente

especificables como fuerza, masa, carga, distancia y velocidad. En un sistema no teleológico la explicación de cualquier proceso desplegado temporalmente tiene que consistir en una explicación por referencia a esas leyes de cómo cada estado del universo se desenvuelve de su inmediato predecesor. La teleología, por el contrario, admitiría principios irreductibles que gobiernan el desarrollo que se despliega temporalmente.

La teleología que estoy considerando sería una explicación no solo de la aparición de los organismos físicos sino del desarrollo de la conciencia y, por último, de la razón en esos organismos. Pero su forma puede describirse incluso si permanecemos en el nivel físico. La teleología natural requeriría dos cosas. Primero, que las leyes no teleológicas y atemporales de la física —esas que gobiernan los elementos últimos del universo físico, cualesquiera que sean— no sean completamente deterministas. Dado el estado físico del universo en un momento dado, las leyes de la física tendrían que posibilitar una gama de estados alternativos consecuentes, presumiblemente con un reparto de probabilidad entre ellos.

Segundo, entre esos futuros posibles habrá algunos que sean más idóneos que otros como pasos posibles en el camino a la formación de sistemas más complejos y, por último, de los tipos de sistemas replicantes característicos de la vida. La existencia de la teleología requiere que los estados sucesivos en este subconjunto tengan una probabilidad significativamente más alta que la que es implicada por las solas leyes de la física, simplemente porque están en camino hacia un cierto resultado. Las leyes teleológicas asignarían una probabilidad más alta a los pasos en un espacio establecido que tiene una más alta «velocidad» hacia ciertos resultados[45]. Serían esencialmente leyes de auto-organización de la materia, o de lo que sea más básico que la materia.

Esta es una hipótesis teleológica franca porque las transiciones preferidas no tienen una probabilidad más alta en virtud de sus características intrínsecas inmediatas, sino solamente en virtud de desarrollos desplegados temporalmente de los cuales forman parte potencial. En otras palabras, algunas leyes de la naturaleza se aplicarían directamente a la relación entre el presente y el futuro más bien que a especificar funciones instantáneas que ocurren todas las veces. Una

teleología naturalista significaría que principios organizacionales y de desarrollo de este tipo son una parte irreductible del orden natural y no el resultado de la influencia intencional o propositiva de nadie. No estoy seguro de que esta idea aristotélica de la teleología sin intención tenga sentido, pero por el momento no veo por qué no.

7

¿Cuáles son las alternativas? Cualquier alternativa debe incluir la posibilidad constitutiva, en la naturaleza de los elementos de los cuales se compone el mundo, de su combinación en organismos vivos con las propiedades de la conciencia, acción y conocimiento que sabemos que tienen. Pero dada esta posibilidad, la cuestión histórica de por qué surgieron tales organismos podría en principio recibir dos respuestas no teleológicas muy diferentes. Primero está la hipótesis de que la aparición inicial del sistema replicante gobernado por código que comenzó el proceso evolutivo fue un accidente cósmico, y que mutaciones accidentales subsiguientes proporcionaron el conjunto de candidatos sucesivos sobre los que la selección natural obró para generar la historia de la vida. Esta hipótesis hace que el resultado sea demasiado accidental para que cuente como una explicación genuina de la existencia de la conciencia, de seres pensantes como tales.

Segundo, para los teístas está la alternativa intencional: la intervención divina crea vida del material básico del mundo y, quizás, también guía el proceso de evolución por selección natural mediante la producción intencional y la preservación de algunas de las mutaciones sobre las que la selección natural obra[46]. Esto podría combinarse o bien con una respuesta reductiva o emergente a la cuestión constitutiva. Una explicación creacionista de la existencia de la vida es el análogo biológico del dualismo en filosofía de la mente. Empuja a la teleología fuera del orden natural en las intenciones de un creador, que trabaja con materiales completamente sin dirección cuyas propiedades a pesar de eso forman la base tanto de lo mental como de lo físico. Si Dios, en algún momento del pasado, construyó ADN o uno de los predecesores de sus elementos, eso excusa la necesidad

de cualquier explicación de la capacidad de los elementos de organizarse ellos mismos de esta aparente forma intencional.

Eso requeriría que la existencia del ADN fuese una posibilidad física, en el espacio químico por decirlo de alguna manera. Y si ampliamos el asunto a la conciencia y a la razón, requeriría que sujetos conscientes y racionales soportados por cerebros del tipo correcto fueran posibilidades mentales. Pero en la imagen creacionista, el orden natural da cuenta de la posibilidad física del ADN de la misma manera que da cuenta de la posibilidad física de un avión, un teléfono o un ordenador. Estas posibilidades se explican todas por la sola física: es solamente su actualización lo que necesita un diseñador; y algo análogo sería verdadero para la conciencia animal, una forma sorprendente en la que los elementos protopsíquicos del mundo pueden combinarse. Así, la organización biológica y mental no es parte del orden natural desde un punto de vista creacionista más de lo que lo son los aviones o los teléfonos. Las leyes de la naturaleza entrañan su posibilidad, pero no explican su ser efectivo.

Mi preferencia por una explicación natural e inmanente es congruente con mi ateísmo. Pero incluso un teísta que cree que Dios es el último responsable de la aparición de la vida consciente podría mantener que esto ocurre como parte de un orden natural que es creado por Dios pero que no requiere de una ulterior intervención divina. Un teísta no comprometido con el dualismo en filosofía de la mente podría suponer que la posibilidad natural de los organismos conscientes reside ya en la naturaleza de los elementos de los que estos organismos están compuestos, quizás complementados por las leyes de la emergencia psicofísica. Hacer de la posibilidad de la vida consciente una consecuencia del orden natural creado por Dios mientras se adscribe su ser de hecho a una intervención divina subsiguiente parecería una complicación arbitraria. Alguna forma de naturalismo teleológico debería, por estas razones, parecer no menos creíble que una explicación intervencionista, incluso a los que creen que Dios es el último responsable de todo[47].

CAPÍTULO 4

EL VALOR

1

La idea de teleología implica alguna clase de valor en el fin hacia el que las cosas tienden, incluso si la teleología se separa de la intención y el fin no es la meta del agente que aspira a ella. Pero permítanme dejar a un lado estas especulaciones cosmológicas y ocuparme de la más mundana existencia y función del valor en nuestras vidas. El valor real —bien y mal, correcto e incorrecto— es otra de esas cosas, como la conciencia y el conocimiento, que parece incompatible con el naturalismo evolucionista en su forma común materialista. En el capítulo anterior comenté brevemente cómo la teoría evolucionista podría reconciliarse con el realismo respecto al tema del valor siguiendo el modelo de una supuesta armonización del realismo científico con la teoría evolucionista. Expresé mis dudas sobre la suficiencia de tal reconciliación en ambos casos, pero ahora quiero tratar el problema del valor y de las razones para la acción con más detalle.

El problema del lugar del valor en el mundo natural incluye, pero trasciende, los problemas de la conciencia y del conocimiento en general porque tiene que ver con el ámbito práctico, el control y la evaluación de la conducta. Resulta claro que la existencia del valor y nuestra respuesta a él dependen de la conciencia y del conocimiento, ya que mucho de lo que resulta valioso consiste en o envuelve experiencia consciente, y las respuestas apropiadas a lo que es bueno y malo, correcto y erróneo, dependen del reconocimiento cognitivo de las cosas que nos da razones a favor y en contra de ellas. He argumentado hasta ahora que la realidad de la

conciencia y el conocimiento no puede ser plausiblemente reconciliada con el naturalismo científico tradicional, tanto constitutiva como históricamente. Creo que el valor presenta un problema ulterior para el naturalismo científico. Incluso en el supuesto de una concepción del mundo en la que se diera a la conciencia y al conocimiento un lugar en el orden natural, el valor es algo añadido y tiene consecuencias que son comparables en alcance.

Para dar razón de este aserto es necesario explicar lo que se entiende por realidad del valor, algo mucho menos transparente que la realidad de la conciencia y también menos claro que la realidad del conocimiento o de la razón en general. Una de las dificultades en esta área es describir una forma de realismo del valor que no esté lastrado con carga metafísica y que sea claramente distinguible de una concepción subjetivista sofisticada del valor. El realismo del valor es altamente controvertido porque su explicación subjetivista no es flagrantemente implausible en la forma en que lo son la explicación subjetivista de la ciencia o la explicación materialista de la conciencia. Así que permítaseme explicar la clase de realismo del valor que presenta problema y cómo hay que distinguirlo de las formas sofisticadas de subjetivismo de la tradición humeana, que opta por versiones antirrealistas de la objetividad moral, la justificación interpersonal y la razón práctica.

En términos simples, la posición subjetivista que contrastaré con el realismo es que la verdad valorativa y moral depende de nuestras disposiciones y respuestas motivacionales mientras que en la posición realista, por el contrario, nuestras respuestas intentan reflejar la verdad valorativa y pueden ser correctas o incorrectas respecto de ella. Para facilitar la economía intelectual y discursiva, voy a tratar el subjetivismo como un punto de vista sobre la verdad y las condiciones de verdad. Pero lo que voy a decir sobre eso y sobre la comparación con el realismo debería aplicarse también a las formas expresivistas igualmente comunes de subjetivismo derivadas de Hume, según las cuales los juicios de valor expresan ciertas clases de actitudes o sentimientos del sujeto más que el ser verdadero o falso en virtud de tales actitudes. Tales puntos de vista pueden normalmente establecerse desde el punto de vista de las condiciones de asertabilidad más que de las condiciones de verdad. El subjetivismo de los

valores y de la moralidad adopta muchas variedades, y espero que sea posible tratar el asunto básico sin distinguir entre ellas.

La diferencia entre subjetivismo y realismo, en lo que se refiere a la dependencia de la verdad de nuestras respuestas, no se aplica a todo tipo de valor. En particular, los realistas pueden estar de acuerdo con los subjetivistas en que el valor de experiencias básicas de placer y dolor, por ejemplo, es inseparable de nuestras respuestas naturales de atracción y aversión a ellas. En estos casos, para un realista, apariencia y realidad coinciden. A la pregunta: «¿Sería el dolor malo incluso si no nos diésemos cuenta de él?». La respuesta es «No» (de hecho no sería ni dolor). Solamente cuando nos trasladamos a la valoración de las experiencias ausentes —las del futuro o las de otros sujetos— o a juicios sobre cómo tratar con posibilidades que envuelven múltiples experiencias, quizás con valores en conflicto, o con juicios sobre el valor de cosas más allá de la experiencia, el realismo y el subjetivismo darán explicaciones claramente divergentes. La posición subjetivista es que la respuesta correcta depende de nuestras actitudes y disposiciones; la posición realista es que nuestros juicios intentan identificar la respuesta correcta y dirigir nuestras actitudes de acuerdo con ella, tanto si la cuestión es acerca del dolor como si lo es sobre cualquier otra cosa[48].

Las formas más plausibles de subjetivismo descansan en algunas variantes de la concepción de Hume de las pasiones, que incluye el sentido moral. Desde este punto de vista los juicios de valor en general y los juicios morales en particular se fundan en aspectos más sofisticados y reflexivos del sistema motivacional que los apetitos básicos y las emociones instintivas. Los juicios prudenciales son la manifestación de una pasión sosegada de auto interés temporalmente imparcial que genera un deseo o aversión igual por el provecho o daño propio presente y futuro. Los juicios morales manifiestan un sentimiento de benevolencia imparcial o, en casos más complejos, un apego a prácticas o instituciones que aumentan el bienestar general o el bien de cada uno. Los pormenores del sentimiento moral pueden ser complicados y estar sujetos a modificación social, pero el punto esencial es que los juicios de valor no son nada más que la expresión de tal sentimiento: se puede decir que son correctos o incorrectos solo por

referencia a él, junto con los hechos no-valorativos que son objeto del juicio. Lo que determina la verdad de un juicio de valor es que resulta de la operación sosegada, reflexiva y completamente informada de estas disposiciones motivacionales.

Una posición realista, por el contrario, niega que la verdad dependa de nuestras disposiciones (del nivel de sentimientos inmediatos como placer y dolor) y sostiene que cuando nuestros juicios de valor son correctos es porque nuestras disposiciones están en consonancia con la estructura de hecho y el peso de los valores en el caso de que se trate. Un juicio que dijera que uno no debe causar un daño importante a otro por buscar un pequeño beneficio propio se fundamenta en el reconocimiento de que la razón contra causar daño es mucho más fuerte que la razón para buscar el beneficio, y que el hecho de que el daño sea sufrido por algún otro no es una razón para desconsiderarlo.

2

¿Pero por qué el interés de otros nos aporta razones para la acción si no estamos predispuestos por reflexión a ser movidos por algún grado de benevolencia? Esa cuestión es la expresión natural del escepticismo del subjetivista sobre el realismo. Es también el primer paso del camino a una interpretación equivocada y fatal del realismo, la malinterpretación que lo convierte en una teoría metafísica. Ello supondría que alguna otra cosa distinta del valor debe hacer valiosos los juicios verdaderos o falsos, algo en el mundo además de los hechos empíricos que cuentan por y contra la acción, tal como el hecho de que si no piso el freno atropellaré a un perro. Los subjetivistas piensan que el realista debe responder la pregunta porque su propia teoría tiene una respuesta para ello: encuentran el fundamento de la verdad de los juicios de valor en hechos psicológicos sobre las disposiciones motivacionales humanas, algo más fundamental que los valores. (Para los expresivistas la cuestión se evita negando que los juicios de valor tengan valor de verdad, pero también encuentran el fundamento de esos juicios en las disposiciones motivacionales.)

Esta clase de explicación del valor es precisamente lo que los realistas niegan. Más importante es, en mi concepción del realismo moral y valorativo, que los realistas no sustituyen la respuesta psicológica a la cuestión planteada por los subjetivistas con una respuesta diferente. El realismo no es una teoría metafísica del fundamento de la verdad moral y valorativa. Es una posición metafísica solamente en el sentido negativo de que niega que toda verdad básica sea solo natural o matemática. Es metafísica solo si la negación de una posición metafísica como el naturalismo también lo es[49]. Pero el realismo del valor no mantiene que los juicios de valor se conviertan en verdaderos o falsos por algo más, ya sea natural o sobrenatural.

Por supuesto que son los hechos naturales los que determinan la verdad de algunos juicios de valor en el sentido en que son hechos que proporcionan razones por y contra la acción[50]. De esta forma, el hecho de que atropellarás a un perro si no pisas los frenos proporciona un argumento para que debas pisar los frenos. Pero la verdad moral general que permite esta inferencia —a saber, que pesa a favor el evitar un doloroso daño a una criatura sentiente— no se hace verdadera por ningún hecho de cualquier otra clase. No es nada sino ella misma.

La disputa entre realismo y subjetivismo no gira en torno a los contenidos del universo. Es una disputa sobre el orden de la explicación normativa. Los realistas creen que la moral y otros juicios valorativos pueden ser explicados por verdades valorativas más generales o básicas junto con los hechos que los traen a escena (el hecho sobre el perro). Pero no creen que el elemento valorativo en tal juicio pueda explicarse por algo más. Que hay una razón para hacer lo que evitará un daño importante a una criatura sentiente es, desde un punto de vista realista, una de la clase de cosas que puede ser verdadera por ella misma y no porque algo de un tipo diferente sea verdadero. En esto recuerda a las verdades físicas, las verdades psicológicas y a las verdades aritméticas y geométricas.

En lugar de explicar la verdad o falsedad de los juicios de valor por su conformidad con nuestras disposiciones motivacionales o con el sentido moral, como hace el subjetivista, el realista explica nuestro sentido moral como una facultad que aspira a identificar los hechos en las circunstancias

de elección que cuentan por y contra ciertas direcciones de la acción y descubrir cómo se combinan para determinar qué rumbo sería el correcto, qué conjunto de alternativas sería permisible o aconsejable y qué otros se descartan. Los hechos que proporcionan estas razones y justificaciones son tan comunes como qué pasaría si uno hiciera esto o aquello, a quién beneficiaría, a quién se dañaría, quién ha prometido qué a quién, etc.

Pero aunque el realista no añada nada al catálogo de entidades o propiedades que un subjetivista cree que existen en el mundo, sostiene de hecho que ciertas verdades, que los subjetivistas piensan que tienen que fundarse en algo más, no tienen que estar fundadas así sino que son verdaderas por propio derecho. Después de todo, cualquiera que sea el punto de vista filosófico de uno, en la medida en que hay tal cosa como la verdad debe haber algunas verdades que no tengan que fundarse en nada más. El desacuerdo sobre qué verdades sean esas define algunas de las líneas más profundamente fallidas de la filosofía. Para los filósofos de una tendencia idealista es obvio que los hechos físicos no pueden ser verdaderos por ellos mismos sino que deben ser explicados desde el punto de vista de la experiencia fáctica o posible, exactamente como es obvio a aquellos de una tendencia materialista que los hechos mentales no pueden ser verdad por ellos mismos y deben explicarse en vistas de la conducta fáctica o posible, la organización funcional o la fisiología. La cuestión sobre el realismo moral es de la misma clase. Alguien que encuentra un reto incontestable en la pregunta: «¿Pero qué es una razón?», muestra una creencia restrictiva del más amplio alcance sobre qué clase de verdad básica pueda haber allí[51].

En todas las áreas de pensamiento debemos confiar por último en nuestros juicios, comprobados por la reflexión, sujetos a corrección por los contra-argumentos de otros, modificados por la imaginación y por la comparación con alternativas. El antirrealismo es siempre una posibilidad conjetural: la pregunta puede siempre ser planteada, si hay algo más que la verdad en un cierto ámbito que nuestra tendencia a alcanzar conclusiones de esta forma, quizás en convergencia con los demás. Algunas veces, como con la gramática o la etiqueta, la respuesta es no. Por esa razón la convicción intuitiva de que en un ámbito particular, como el mundo físico o

las matemáticas o la moralidad o la estética, nuestros juicios son intentos de responder a una clase de verdad que es independiente de ellos, puede ser imposible de establecer definitivamente. Y eso sin perjuicio de que de todas formas sea muy sólida y justificada.

Hay explicaciones subjetivistas alternativas de la aparición de la independencia de la mente en la verdad de la moral y otros juicios de valor. Una de las cosas que un subjetivismo sofisticado nos permite decir cuando juzgamos que el infanticidio es erróneo es que sería erróneo incluso si ninguno de nosotros pensara así, aunque ese segundo juicio también está fundado en nuestras respuestas. Sin embargo, encuentro tan implausibles estas explicaciones cuasi-realistas, expresivistas, del fundamento de la objetividad de los juicios morales como la explicación subjetivista de los juicios de valor más simples. Estos epiciclos son de la misma clase que la propuesta original: niegan que los juicios de valor puedan ser verdad por propio derecho y esto no concuerda con la que creo que es la mejor explicación general de nuestro pensamiento sobre el valor.

No hay un experimento crucial que establecerá o refutará el realismo del valor. Un fundamento para rechazarlo, el tipo usado por Hume, es simplemente una petición de principios: si se supone que las verdades morales objetivas pueden existir solamente si son como otra clase de hechos —físicos, psicológicos o lógicos— entonces resulta claro que no las hay. Pero el fallo de este argumento es que no prueba que *hay* verdades morales objetivas. Un apoyo positivo para el realismo puede venir solamente de la fecundidad del pensamiento valorativo y moral para producir resultados, incluyendo correcciones de creencias antiguamente sostenidas por todo el mundo y el desarrollo de nuevos y perfeccionados métodos y argumentos conforme avanza el tiempo. La interpretación realista de lo que estamos haciendo al pensar en estas cosas puede convencer solamente si es una explicación mejor que la alternativa subjetivista o socio-constructivista. Siempre va a ser una cuestión de comparación y un asunto de juicio, como lo es en cualquier otro ámbito, ya sea matemáticas o ciencia o historia o estética.

3

Todo esto es un preludio a una pregunta de mayor alcance: ¿cuáles son las implicaciones para el orden natural de las diferentes concepciones del valor? Mi insistencia en que el realismo del valor no es un postulado metafísico de entidades o propiedades extra pudiera sugerir que el realismo no tiene implicaciones en el orden natural. Pero no es así. En esencia, estoy de acuerdo con la postura de Sharon Street de que el realismo moral es incompatible con la explicación darwinista de la influencia evolutiva en nuestras facultades del juicio moral y valorativo[52]. Street sostiene que la explicación darwinista está firmemente apoyada por la ciencia contemporánea, por lo que concluye que el realismo moral es falso. Sigo la misma inferencia en la dirección opuesta: en tanto que el realismo moral es verdadero, una explicación darwinista de los motivos subyacentes al juicio moral debe ser falsa a pesar del consenso científico en su favor.

Pero las implicaciones son mucho más amplias. Ya que somos, evidentemente, productos de la evolución y en última instancia de un proceso cósmico que llevó al desarrollo primero de organismos unicelulares y después de agentes conscientes antes de producir por último seres inteligentes capaces de juicios de valor, la imagen del orden natural que hace posible este proceso debe ser ampliada. Una concepción adecuada del cosmos debe contener los recursos para explicar cómo pueden haber surgido seres capaces de pensar exitosamente sobre lo que es bueno y malo, correcto e incorrecto, y descubrir verdades morales y valorativas que no dependen de sus propias creencias. Esto es análogo a las implicaciones previamente defendidas para el orden natural de la existencia de la conciencia y el conocimiento, pero las trasciende.

Se han hecho comentarios sobre la extraña naturaleza, rompe-categorías, de esta disputa. El argumento de Street se apoya en una aserción científica empírica para refutar una postura filosófica en metaética. Yo, incluso más extrañamente aún, me apoyo en un aserto filosófico para refutar una teoría científica apoyada por la evidencia empírica. Pero no creo que sea inadecuado que se genere pensamiento sobre la cuestión. Los juicios de valor y los razonamientos morales son parte de la vida humana y,

por tanto, parten de la evidencia fáctica de la que los humanos son capaces. La interpretación de facultades tales como estas es inevitablemente relevante para la tarea de descubrir la mejor explicación científica o cosmológica de lo que somos nosotros y cómo llegamos a la existencia. Lo que vale como una buena explicación depende con fuerza de una comprensión de qué es lo que tiene que ser explicado.

Permítaseme comenzar con un bosquejo del argumento de Street de que el realismo moral es incompatible con el darwinismo. Se encuentra en su artículo «Un dilema darwinista para las teorías realistas del valor»[53], que es bastante intrincado y comprende más de una forma de realismo. Pero el argumento se aplica a lo que ella llama versiones *no naturalistas* del realismo del valor, según las cuales los hechos o verdades valorativas no son reductibles a ningún tipo de hecho natural y no requieren misteriosas propiedades «extra» en el mundo, pero son hechos o verdades normativas irreductibles. Es la clase de realismo que he descrito arriba.

Street señala que, si las respuestas y facultades que generan nuestros juicios de valor son en una parte significativa resultado de la selección natural, no hay razón para esperar que nos capaciten para detectar ninguna verdad moral o valorativa independiente de la mente, si es que hay algo así. Porque la capacidad para detectar tal verdad, a diferencia de la capacidad de detectar la verdad independientemente de la mente sobre el mundo físico, no aportaría nada a la eficacia reproductiva. No es del todo implausible que las disposiciones motivacionales valorativas características de los seres humanos, y sus capacidades para el razonamiento práctico y para la convergencia interpersonal en las prácticas y formas de justificación, puedan deberse en una medida importante, ya sea directa o indirectamente, a la selección natural darwinista. Pero para esta explicación es completamente irrelevante si esas facultades nos capacitan para detectar verdad moral independiente de la mente, si hubiera tal cosa, o si realmente nos extravían.

Street observa que la explicación natural darwinista de los motivos y disposiciones que forman los puntos de partida de nuestros juicios de valor, y que podemos modificar mediante el proceso del equilibrio reflexivo, es que han contribuido a la eficacia reproductiva no solo ayudando a la

supervivencia individual sino promoviendo la crianza y el cuidado de los niños, impidiendo la agresión y haciendo posible la cooperación social. La verdad independiente de la mente de los juicios resultantes no desempeña ninguna función en la historia darwinista: en lo que a la selección natural concierne, si hubiera tal cosa como una verdad moral independiente de la mente, estos juicios podrían ser sistemáticamente falsos.

No puede decirse lo mismo de nuestros juicios fácticos. Si hay un mundo físico independiente de la mente, la incapacidad sistemática para detectar la verdad básica en nuestro entorno (dejando de lado una verdad científica más sofisticada) sería desastrosa para nuestra eficacia reproductiva. El realismo sobre el mundo físico es un aspecto fundamental de cualquier explicación darwinista de nuestras facultades perceptivas y cognoscitivas, tanto como de nuestros motivos y capacidades para la acción. Pero el realismo del valor es irrelevante.

Esto requiere que vuelva a algo que dije en el capítulo anterior. Allí ofrecí una forma en la que el realismo moral podría combinarse con una explicación darwinista de nuestras facultades cognoscitivas, incluyendo la facultad de la razón práctica. Dije que si podemos comprender nuestras impresiones prerreflexivas del valor —atracciones y aversiones instintivas, inclinaciones e inhibiciones— como apariencias del valor real, entonces el proceso cognitivo de descubrir una estructura sistemática y consistente de razones generales, prácticas y principios morales puede pensarse como una forma de ir de la apariencia a la realidad en el ámbito normativo. A ese respecto sería análogo a la explicación darwinista de la razón científica, construida realistamente. Después continué estableciendo un argumento de un tipo diferente para no aceptar una explicación darwinista de la razón; pero permítanme dejar a un lado eso por el momento, porque quiero concentrarme en un problema específico para el realismo del valor que se debe a una diferencia entre los datos prerracionales con los que comienzan el razonamiento fáctico y el razonamiento práctico. Por comodidad limitaré la argumentación al placer y al dolor, aunque hay otros ejemplos.

La gran diferencia entre los dos casos desde un punto de vista darwinista se da entre las apariencias prerreflexivas del valor y las apariencias prerreflexivas del mundo físico[54]. Una concepción realista del

último, por ejemplo de la percepción visual, es esencial a la historia darwinista: la visión contribuye a la eficacia reproductiva porque nos capacita para ver lo de fuera, que es necesario para cualquier tipo de funcionamiento exitoso. Por el contrario, la maldad real del dolor y la capacidad para reconocer esa maldad son completamente superfluas en una explicación darwinista de nuestra aversión al dolor. La aversión al dolor afianza la eficacia solamente en virtud del hecho de que nos conduce a evitar el daño asociado al dolor, no en virtud del hecho de que el dolor es realmente malo. En lo que concierne a la selección natural, el dolor podría perfectamente ser bueno en sí mismo y el placer en sí mismo malo, o (más probablemente) ambos podrían carecer de valor en sí mismos, aunque naturalmente seamos ciegos a ese hecho.

Un realista, a diferencia de un subjetivista, no puede ver estas posibilidades como carentes de significado, sino que debe en su lugar creer que son falsas, exactamente como un realista del mundo físico cree que la hipótesis del genio maligno sobre la percepción visual más que un sin sentido es falsa. Una explicación darwinista de la percepción visual implica que nos da información acerca del mundo externo y que la hipótesis del genio maligno es falsa. Una explicación darwinista del origen de nuestros deseos y aversiones básicas, por el contrario, no tiene implicaciones sobre si generalmente son percepciones fiables de valores independientes del juicio o si en verdad hay tales cosas.

4

Así que estoy de acuerdo con Street en que, desde una perspectiva darwinista, la hipótesis del realismo del valor es superflua, una rueda que gira sin estar enganchada a nada. Desde una perspectiva darwinista nuestras impresiones de valor, si están construidas realistamente, son completamente infundadas. Y si eso es verdad para nuestras respuestas más básicas, es también verdad para toda la estructura del valor y la moralidad que se construye sobre ellas por la reflexión práctica y el desarrollo cultural; exactamente como un realismo científico se destruiría si abandonamos una interpretación realista de las experiencias perceptuales en las que la ciencia

se funda. Un sistema basado en mantener la coherencia o la consistencia de las respuestas propias no necesita una idea de verdad independiente de la mente respecto del valor (como opuesto a la lógica) si las respuestas que proporciona su contenido original no se refieren a nada más allá de ellas mismas.

A pesar de ello estoy convencido de que el dolor es realmente malo y no solo algo que odiamos, y de que el placer es realmente bueno y no solo algo que nos gusta. Eso es manifiesta y exactamente como me parecen, por mucho que intente imaginar lo contrario, y sospecho que lo mismo le pasa a la mayoría de la gente. Eso no significa que nuestras respuestas viscerales sean infalibles, no más de lo que nuestras percepciones visuales prerreflexivas lo son. Ellas son simplemente los puntos de partida para explorar un ámbito que puede requerir un extenso razonamiento práctico y moral para comprenderlo.

En la explicación darwinista debe considerarse como una ilusión, quizás una ilusión de objetividad que es producto de la selección natural por su contribución a la eficacia reproductiva. En verdad, la disposición para adscribir una objetividad ilusoria a normas claramente contingentes dependientes de respuesta, de lenguaje y costumbre por ejemplo, parece ser bastante típica de los humanos, y bastante útil. Sin embargo, en mi caso, las credenciales científicas del darwinismo y estos otros ejemplos, no bastan para eliminar mi convicción inmediata de que la objetividad no es una ilusión en lo que se refiere a los juicios básicos de valor.

¿Pero cuál es la alternativa realista? Describirla es complicado ya que es obvio que, estando conectados biológicamente placer y dolor, desempeñan biológicamente una función vital en la eficacia reproductiva de las criaturas conscientes, incluso si su valor objetivo no lo hace. La posición realista debe ser que estas experiencias que incorporan deseo y aversión como parte de su esencia también tienen valor positivo y negativo en ellas mismas, y que nos es evidente por reflexión, aunque no forman parte necesaria de la explicación evolutiva de por qué están asociados con ciertos episodios corporales tales como el sexo, comer o herirse. Son adaptativos, pero son algo más que eso. Aunque no son las únicas cosas que tienen valor objetivo, estas experiencias están entre los fenómenos más destacados por los que el

valor entra en el universo y son los ejemplos más claros por los que llegamos a familiarizarnos con el valor real.

En la interpretación realista, el placer y el dolor tienen una doble naturaleza. En virtud de la atracción y aversión que le son esenciales desempeñan una función vital en la supervivencia y la eficacia adaptativa, y su asociación con funciones y malfuncionamientos biológicos específicos puede ser explicada por selección natural. Pero para seres como nosotros, capaces de razón práctica, son también objetos de conciencia reflexiva, comenzando con el juicio de que el placer y el dolor son buenos y malos por ellos mismos y nos conducen, junto con otros valores, al reconocimiento más sistemático y elaborado de razones para la acción y principios que gobiernan su combinación e interacción y, en último lugar, a los principios morales.

Si a las facultades que generan estos juicios se les pudiera también dar una explicación darwinista, entonces la interpretación realista sería refutada porque no tendríamos razón para considerarlas como descubrimientos de lo que es verdadero independientemente de nuestros juicios. La bondad objetiva del placer y la maldad del dolor, y la verdad o falsedad objetiva de nuestros más complejos juicios de valor, serían completamente irrelevantes para la comprensión de estas facultades o para lo que estamos haciendo cuando las ejercitamos. Por tanto desde el punto de vista del realista, aunque la asociación de placer y dolor con sexo y herida pueda ser explicada por selección natural, sus valores objetivos, nuestra capacidad para reconocerlos y todo lo que sigue de esto, no pueden serlo.

Street concluye que el realismo no puede estar en lo cierto, yo concluyo que algo se echa en falta en el darwinismo y en la concepción biológica estándar de nosotros mismos. La conclusión negativa de Street viene con una alternativa positiva, una forma de constructivismo[55]. Desafortunadamente mi conclusión negativa deja una alternativa positiva bastante misteriosa y, a menos que sea posible decir más, será difícil comparar cómo responden los puntos de vista en liza a la premisa hipotética compartida.

5

Como con la situación de la conciencia y el conocimiento en el orden natural, el problema del valor y el orden natural tiene tanto un aspecto constitutivo como otro histórico. La pregunta constitutiva gira sobre nuestra naturaleza: ¿Qué clase de seres somos si el realismo es verdadero y en verdad reconocemos y respondemos a valores y razones prácticas que no son solo producto de nuestras propias respuestas? La pregunta histórica versa sobre nuestros orígenes: ¿Cómo debe ser el universo y el proceso evolutivo para haber generado tales seres? Ambas cuestiones parecen requerir alguna alternativa al naturalismo materialista y su aplicación darwinista en biología, pero ¿cuáles son las posibilidades?

La consecuencia constitutiva más destacada del realismo sería que los seres humanos pueden no solo detectar sino también ser movidos por el valor. En el caso de valores básicos de la experiencia tales como la bondad del placer y la maldad del dolor, una motivación instintiva nace de la experiencia misma: el deseo de que continúe es parte del placer y el deseo de que pare es parte del dolor. Pero no ocurre eso cuando pensamos en el placer o dolor que no está presente de hecho —ya sea en nuestra experiencia de otras ocasiones o en la experiencia de otros—. Incluso podemos dirigirnos hacia él, aún reconociendo que el dolor es malo y que hay razones para evitarlo tanto para nosotros mismos como para otros. Tales consideraciones pueden llevarnos a resistir lo inmediato, motivación inherente del placer o dolor presentes, dejando solamente su valor objetivo. Y por supuesto lo mismo es verdad de valores mucho más complejos tales como la honestidad y la deshonestidad, la justicia y la injusticia, la lealtad y la traición.

¿Qué es esta capacidad motivacional? Como he dicho, no debería verse como un componente metafísico del mundo extra que ejerce una influencia causal sobre nosotros. Los rasgos del mundo que confieren valor y proporcionan razones son de ordinario hechos de las experiencias de la gente, sus relaciones de unos con otros, y las implicaciones para sus vidas y las de otras criaturas en diferentes vertientes de la acción. Una razón para la acción es un hecho ordinario, tal como el hecho de que la aspirina te curará

el dolor de cabeza, y ser una razón pesa a favor de tomarte la aspirina o de que te dé una. Mi capacidad para responder a los valores reales es la capacidad, por ejemplo, de motivarme a darte una aspirina por esa razón, porque sé que te curará el dolor de cabeza y reconozco que eso pesa a favor porque los dolores de cabeza son malos. Hacemos esta clase de cosas todo el tiempo. ¿Qué significa?

Creo que envuelve un control consciente de la acción, que no puede ser analizado como causalidad física con un acompañamiento consciente epifenoménico, y que incluye alguna forma de voluntad libre, aunque es, como siempre, muy obscuro qué sentido dar a esa noción. Respondo conscientemente al valor cuando decido darte una aspirina porque aprendo que tienes dolor de cabeza y sé que la aspirina te pondrá mejor. Por supuesto quiero que tu dolor de cabeza se cure, pero eso es también el resultado de mi reconocimiento de que los dolores de cabeza son malos. La explicación de mi acción se refiere a estos hechos acerca de los dolores de cabeza y a la aspirina en su estatus de razones, como contando en pro y en contra de hacer ciertas cosas. Afectan a nuestra conducta porque son reconocidas como razones por un agente sensible al valor. No es tan diferente de la forma en la que el reconocimiento de razones en un argumento puede explicar la formación de una creencia fáctica, como expuse en el último capítulo.

Mis convicciones sobre el problema tradicional de la voluntad libre son incompatibilistas, pero no estoy seguro de que ese asunto tenga que dirimirse para los propósitos presentes. Si hay una forma en que la motivación consciente por razones para la acción pueda ser reconciliada con el determinismo causal de la acción, ya sea físico o psicológico, entonces el realismo del valor puede ser compatible con el determinismo. De momento, solamente deseo insistir en que no es compatible con una concepción darwinista de cómo se constituyen los orígenes de nuestros motivos.

Si dejamos a un lado el asunto del determinismo, la concepción distintiva de los seres humanos inherente al realismo del valor es que pueden ser motivados por su aprehensión de valores y razones. Su existencia es un tipo básico de verdad y la explicación de la acción por tales

motivos es una forma básica de explicación no reductible a ninguna otra forma, ya sea psicológica o física. Te doy una aspirina porque sé que te aliviará el dolor de cabeza; con ello manifiesto mi reconocimiento de que este hecho pesa a favor de la acción. Y esto no es simplemente una descripción superficial de algo más, que es lo real: la explicación subyacente.

La acción humana, en otras palabras, se explica no solo por la fisiología o por los deseos, sino por los juicios. Somos sujetos de actitudes sensibles de juicio, en expresión de Scanlon[56], y estos juicios tienen un objeto más allá de ellos mismos. Existimos en un mundo de valores y respondemos a ellos a través de juicios normativos que guían nuestras acciones. Esto, como nuestras capacidades cognitivas más generales, es un desarrollo más alto de nuestra naturaleza como criaturas conscientes. Quizás incluya también la capacidad de responder al valor estético, realistamente construido, como un ámbito de juicio independiente que nos revelan nuestras experiencias y juicios.

Si incidimos de nuevo sobre la elección entre explicaciones reductivas y emergentes, no parece probable que el ejercicio de la conciencia en los juicios valorativos y el razonamiento práctico nos conduzca a un análisis reductivo incluso si un monismo reductivo explica la existencia de la conciencia en los organismos complejos. Como con el conocimiento en general, la respuesta al valor parece solamente tener sentido como una función del sujeto unificado de la conciencia y no como una combinación de las respuestas de sus partes. En ese respecto es diferente de las experiencias de placer y dolor que proporcionan algunas de las más importantes materias primas del valor: podrían explicarse reductivamente mediante alguna forma de monismo psicofísico. Pero el razonamiento práctico y su influencia en la acción envuelve al sujeto consciente unificado que ve lo que debería hacer y que sugiere una respuesta emergente a la pregunta constitutiva. Por otro lado, el problema de cómo reconciliar la unidad del sujeto, tal como es, con una explicación reductiva de lo mental aqueja al monismo psicofísico a todos los niveles. Quizás en este escenario extravagantemente especulativo el problema parte-todo no debería excluir

el reductivismo de la razón práctica tan decisivamente como lo hace con las sensaciones de gusto.

El aspecto metafísico más importante de un punto de vista realista de la razón práctica es que la conciencia no es epifenoménica y pasiva sino que desempeña una función activa en el mundo. Esto es parte del punto de vista ordinario que tenemos de nosotros mismos, pero añade a la irreductibilidad psicofísica de la conciencia empírica un aspecto ulterior en el que la forma materialista de naturalismo fracasa en proporcionar una explicación completa de la naturaleza de los seres humanos[57]. Si la razón práctica es emergente o reductible a la actividad en el micro nivel mediante alguna forma de monismo psicofísico, el realismo del valor requiere que la conciencia sea activa y excluir el epifenomenalismo de la acción humana. (Un punto de vista subjetivista del valor podría ser también incompatible con el epifenomenalismo, pero no tiene por qué: el subjetivismo podría ser consistente con la posibilidad de que nuestros juicios de valor conscientes sean todos efectos colaterales de las causas fisiológicas de la acción.)

Dejando este asunto aparte, volvamos a la diferencia esencial entre una concepción subjetiva y una realista del valor. El subjetivismo interpreta nuestros juicios de valor como un brote y elaboración de nuestras disposiciones motivacionales naturales, hecho posible nada más que por nuestras capacidades lingüísticas y racionales. El realismo las interpreta como el resultado de un proceso de descubrimiento, comenzando desde las apariciones iniciales del valor que son comparables a las creencias perceptivas y trasladándonos (esperamos) hacia una mejor comprensión de cómo deberíamos vivir. Si el realismo es verdadero, la razón práctica es una de nuestras facultades cognoscitivas.

6

No he indicado la respuesta del realismo a la pregunta constitutiva, solo he indicado las condiciones que tal respuesta tendría que satisfacer. Permítaseme ahora decir algo, igualmente poco concluyente, sobre la pregunta histórica. Está claro que, a diferencia del realismo, el subjetivismo apoya, al menos especulativamente, la explicación darwinista de cómo han

surgido las criaturas capaces de tener valores. Al menos esta cuestión no añade nada a las dificultades de tal explicación ya expuestas respecto de los problemas de la conciencia y de la razón en general. El valor comienza, desde este punto de vista, en nuestros deseos e inclinaciones, que son hechos naturales de la psicología animal y humana, y los juicios de valor de más alto nivel son elaboraciones motivacionales sobre esta base, generados por la experiencia, la reflexión y la cultura.

En un punto de vista realista, por el contrario, la pregunta histórica es mucho más oscura, en parte porque el resultado es oscuro. Tenemos que saber lo que tiene que añadirse a la imagen darwinista estándar para que explique la aparición por evolución de criaturas como nosotros, que pueden controlar sus acciones en respuesta a razones. Esta capacidad debe ser consistente con la influencia de la selección natural en que no sea adversa a la eficacia reproductiva, y por tanto no es probable que se extinga por selección natural. Pero su aparición en el menú evolucionista tendría que explicarse por algo más.

El realismo del valor debe dar sentido al hecho de que el proceso biológico evolutivo y la histórica física y química que le precedió han dado lugar a criaturas conscientes, al valor real que llena sus vidas y experiencias y, por último, a seres auto-conscientes capaces de actitudes sensibles de juicio que pueden responder a y ser racionalmente motivados por su conciencia de esos valores. La historia incluye enormes cantidades de dolor tanto como de placer, así que no se presta a una interpretación teleológica optimista. A pesar de eso, el desarrollo del valor y la comprensión moral, como el desarrollo del conocimiento y la razón, y el desarrollo de la conciencia que están en la base de ambas funciones de más alto orden, forman parte de lo que debe explicar una concepción general del cosmos. Como he dicho, el proceso parece ser como el despertarse gradual del universo.

¿Cuál es de hecho la historia del valor en el mundo en la medida en que somos conscientes de ella? Nada puede considerarse obvio en este ámbito, pero en su sentido más amplio, parece coincidir con la historia de la vida. Primero, con la aparición de la vida en sus formas más tempranas, vienen a la existencia entidades que *tienen* un bien, y por el cual las cosas pueden ir

bien o mal. Incluso una bacteria tiene un bien en este sentido, en virtud de su propio funcionamiento, mientras una roca no lo tiene. Por último en el curso de la historia evolutiva aparecen seres conscientes, cuyas vidas empíricas pueden ir bien o mal de formas que nos son conocidas. Más tarde algunos descendientes de estos seres, capaces de reflexión y auto-conciencia, llegan a reconocer lo que les pasa como bueno y malo, y a reconocer razones para buscar o evitar estas cosas. Aprenden a pensar sobre cómo se combinan estas razones para determinar lo que deberían hacer. Y finalmente desarrollan la capacidad colectiva de pensar sobre razones que pueden tener y que no dependen solo de lo que es bueno o malo para ellos mismos.

Esto comienza con las vidas de otros seres como ellos mismos, pero la pregunta puede extenderse al bien y al mal allá donde se encuentren, en las vidas de otras criaturas conscientes y, quizás, incluso en formas de vida faltas de experiencia consciente. Los motivos para la acción se aplican solamente a los seres con razón, y el valor puede entenderse como la explicación de los motivos que tenemos, pero el concepto de valor tiene un espectro de aplicación más amplio que ese. Solo seres capaces de razón práctica pueden reconocer el valor, pero una vez que lo reconocen, lo encuentran en las vidas de las criaturas sin razón práctica. En el más amplio sentido es probablemente coextensivo con la vida, aunque de cuánto de este valor tenemos los humanos motivo para ocuparnos es una cuestión que dejaré abierta. Parece demasiado simplista sostener que el valor genera razones que afectan solamente a las vidas conscientes. Como Scanlon dice, sería cruel y objetable cortar un árbol centenario solo por la diversión de probar la cadena nueva de nuestra motosierra[58].

Para responder a la pregunta histórica sobre el valor y lo valioso se deben responder dos cosas: primero, la aparición del valor en la miríada de formas que adopta en la variedad de vidas capaces de tener un bien; y, segundo, la aparición de motivos para la acción y de los seres capaces de reconocerlos y actuar por ellos. De la primera daría cuenta cualquiera que explique la existencia de la vida, incluyendo el lugar de la conciencia en sus múltiples formas en la vida animal. La segunda requiere algo más. ¿Requiere más que una explicación de la aparición de la razón en un sentido

general? Creo que sí, porque lo que tiene que explicarse es la aparición de la capacidad de reconocer y responder a razones para la acción y no solo una capacidad general cognoscitiva. La razón práctica es un desarrollo del sistema motivacional y de la voluntad, no simplemente un desarrollo del sistema para formar creencias.

Si el valor está ligado a la vida, su contenido dependerá de las formas particulares de vida, y las razones más destacadas que tenemos dependerán, incluso en una concepción realista, de nuestra propia forma de vida. Así es cómo una explicación realista puede incorporar una de las cosas que hacen parecer al subjetivismo máximamente plausible, a saber, el hecho de que lo que encontramos valioso es la mayor parte de las veces una contingencia de lo específicamente biológico de nuestra forma de vida. El bien y el mal humano dependen en primera instancia de nuestros apetitos naturales, emociones, capacidades y lazos interpersonales. Si nos pareciésemos más a las abejas o a los leones lo que nos resulta bueno sería diferente, un punto en el que insiste Street[59].

Sería erróneo intentar encontrar un denominador común, como el placer y el dolor, para integrar en una sola concepción realista los diversos valores que son generados por todas las formas de vida existentes, por no mencionar las imaginables. En su lugar, el valor debe verse como plural: el ámbito del valor real, si hay tal cosa, es tan rico y complejo como la variedad de las formas de la vida, o al menos de la vida consciente. De la misma forma que la mayoría de esas vidas solo son vagamente accesibles desde el interior a nuestro entendimiento, así el valor que generan, positivo y negativo, desborda ampliamente nuestra comprensión. Tampoco resultan claras cuántas razones diferentes pueden generar estos valores, como es obvio por los valores que reconocemos en formas de vida humana diferentes de la propia. ¿Y quién sabe qué formas inimaginables de vida y sus valores asociados existen en otras partes del universo y que no están relacionadas con nosotros por un ancestro común? Pero en tanto que el realismo del valor puede incorporar razones relativas al agente para la acción, el reconocimiento de qué es objetivamente valioso en la vida de una criatura no establece automáticamente la cuestión de qué razones implica para las acciones de otra.

Sin embargo, nuestro acceso directo al valor proviene de la vida humana, la vida de un tipo altamente específico de organismo en la cultura específica que ha creado. El mundo humano, o cualquier vida humana individual, es potencialmente, y frecuentemente de hecho, escena de increíbles riquezas: belleza, amor, placer, conocimiento y la pura alegría de existir y vivir en el mundo. Es también potencialmente, y frecuentemente de hecho, escena de una horrible miseria, pero en ambas situaciones el valor, por específico que pueda ser a nuestra forma de vida, parece ineludiblemente real. Nuestra predisposición a muchos de esos bienes y males desempeña una función vital en nuestra supervivencia y eficacia reproductiva —placer sexual, dolor físico, las punzadas y satisfacciones del hambre y la sed— pero son también buenas y malas en sí mismas y podemos reconocer y sopesar esos valores. Inicialmente los reconocemos en nuestras propias vidas, pero no pueden quedar ahí.

Buscando una explicación histórica, un realista debe suponer que los aspectos más estimulantes de la vida y la conciencia aparecen ya transidos de valor, aunque encuentren su lugar en el mundo a través de sus funciones en las vidas de los organismos que son sus sujetos. Los placeres del sexo, la comida y la bebida son maravillosos, además de ser adaptativos. El valor entra en el mundo con la vida y la capacidad para reconocer el valor y ser influenciado por él en su más amplia extensión aparece con las formas de vida más elevadas. Por tanto la explicación histórica de la vida debe incluir una explicación del valor exactamente como debe incluir una explicación de la conciencia.

Si recordamos los tres tipos posibles de explicación histórica —causal, teleológica e intencional— es difícil justificar una explicación causal. Incluso si hubiera una respuesta en parte reductiva a la cuestión constitutiva de la existencia del valor —si el valor de una experiencia de placer fuese constituido, por ejemplo, por el valor combinado de sus partes protomentales— eso no lleva a ninguna parte en lo que se refiere a la cuestión histórica. Es difícil imaginar qué forma de monismo psicofísico hace posible una explicación reductiva histórica del origen de la vida, el desarrollo de la vida consciente y la aparición de la razón práctica que

convirtiera en algo más que en un mero accidente aquello que consideramos que tiene un valor objetivo.

Por el contrario, una vez que reconocemos que una explicación de la aparición y desarrollo de la vida debe al mismo tiempo ser una explicación de la aparición y desarrollo del valor, es más idónea una explicación teleológica. Esto significaría que lo que explica la aparición de la vida es, en parte, el hecho de que la vida es una condición necesaria de la aparición del valor y, en último lugar, de su reconocimiento.

Dejaré aparte de nuevo la hipótesis de una explicación intencional aunque, también, podría cumplir esta condición. Nos queda entonces la teleología. Según la hipótesis de la teleología natural, el mundo natural tendría una propensión a dar lugar a seres para los que las cosas pueden ser buenas o malas[60]. Esas son todas las formas de vida, las de hecho y las posibles. Han aparecido mediante el proceso histórico de la evolución, pero parte de la explicación de la existencia de ese proceso y de las posibilidades sobre las que opera la selección natural sería que ellas trajeran el valor al mundo, en una gran variedad de formas.

En tanto que la emergencia del valor es la emergencia tanto del bien como del mal, no es candidato para una explicación teleológica puramente benévola: una tendencia hacia el bien. De hecho, ningún principio teleológico que tienda hacia la producción de un solo resultado parece conveniente. Más bien, tendría que ser una tendencia hacia la proliferación de formas complejas y la generación de variaciones múltiples en la gama de los sistemas posibles complejos[61].

Si no estuviéramos inclinados a reconocer razones objetivas para la acción y estuviésemos motivados exclusivamente por nuestros deseos, no tendríamos razón para creer en la existencia del valor en un sentido realista. No habría nada que explicar más allá de motivos subjetivos y la capacidad para ser guiados por la información recibida por la percepción, la memoria y la razón teórica. Pero si concedemos que nuestras impresiones sobre el valor objetivo son sustancialmente correctas, más que completamente ilusorias, entonces debemos considerar la aparición y evolución de la vida como algo más que una historia del desarrollo de organismos que se auto-reproducen, como sostiene la versión darwinista.

Reconocemos que la evolución ha dado lugar a múltiples organismos que *tienen* un bien de tal forma que las cosas les pueden ir bien o mal y que en algunos de esos organismos ha aparecido la capacidad adicional de aspirar conscientemente a su propio bien y, por último, a lo que es bueno en sí mismo. Desde una perspectiva realista esto no puede concebirse simplemente como un efecto accidental y colateral de la selección natural. Una explicación teleológica satisface esta condición. En una explicación teleológica la existencia del valor no es un accidente, porque es parte de la explicación de por qué hay algo como la vida, con todas sus posibilidades de desarrollo y variación. En pocas palabras, el valor no es solo un efecto accidental y colateral de la vida, hay vida porque la vida es una condición necesaria del valor[62].

Esta es una revisión de la imagen darwinista y no una negación completa de ella. Una hipótesis teleológica reconocerá que los pormenores del desarrollo histórico se explican por selección natural entre las posibilidades disponibles sobre la base de la eficacia reproductiva en medios ambientes cambiantes. Pero aunque la selección natural establezca en parte las diferentes formas de vida y conciencia que existen además de las relaciones entre ellas, la existencia del material genético y las formas posibles que lo hacen disponible para la selección tienen que ser explicados de alguna otra forma. La hipótesis teleológica consiste en que estas cosas pueden determinarse no meramente por una química y una física libres de valor sino también por algo más, a saber, una predisposición cósmica a la formación de la vida, la conciencia y el valor que es inseparable de ellas.

En el presente clima intelectual es improbable que tal posibilidad se considere seriamente, pero repetiría mi observación de que no hay disponible ninguna explicación viable, siquiera especulativa, de cómo un sistema tan increíblemente complejo funcionalmente y rico en información como una célula auto-replicante, controlada por ADN, ARN o algún predecesor, podría haber surgido por la sola evolución química de un medio ambiente muerto. Reconocer el problema no se limita a los defensores del diseño inteligente. Aunque los científicos continúan buscando una explicación puramente química del origen de la vida, hay también científicos naturalistas de renombre como Francis Crick que dicen que

parece casi un milagro[63]. Crick se guía por su reflexión sobre las probabilidades de la hipótesis de la «panspermia dirigida», que la Tierra fue sembrada de vida unicelular enviada desde una civilización avanzada en algún lugar de nuestra galaxia donde la vida se había desarrollado antes. Esto depende de la suposición de que habría otros planetas de otras estrellas cuyos medios ambientes físicos harían menos improbable la formación accidental de vida. Pero Crick reconoce que no hay base para confiar en ninguna de estas alternativas.

Alguna forma de teleología natural, un tipo de explicación cuya inteligibilidad defendí brevemente en el último capítulo, sería una alternativa al milagro, ya sea en el sentido de una causalidad salvajemente improbable o en el sentido de una intervención divina en el orden natural. La tendencia a formarse de la vida puede ser un rasgo básico del orden natural que no es explicado por las leyes no-teleológicas de la física y la química. Esto parece una conjetura admisible dada la evidencia disponible. Y una vez que hay seres que pueden responder al valor, una teleología de la acción intencional bastante diferente se convierte en parte de la imagen histórica, que acaba en la creación de nuevo valor. El universo ha llegado a ser no solamente consciente y sabedor de sí mismo sino también capaz, de alguna forma, de elegir su camino hacia el futuro aunque todas estas tres cosas, la conciencia, el conocimiento y la elección, están dispersas en una vasta multitud de seres que actúan tanto individual como colectivamente.

Estas especulaciones teleológicas se ofrecen simplemente como posibilidades, sin convicción positiva. De lo que estoy convencido es de la necesidad de que, para comprender realistamente nuestros problemas y juicios sobre los valores y las razones, debemos rechazar la idea de que resultan de la operación de facultades que se han formado partiendo de cero, por casualidad más la selección natural, o que son efectos colaterales fortuitos de la selección natural o productos de la deriva genética. Cuando nos preguntamos, por ejemplo, si la venganza es una justificación verdadera o solo un motivo natural, o qué tipo de peso deberíamos dar a los intereses de los extranjeros o a los de otras especies, debemos pensar en nosotros mismos como recurriendo a una capacidad de juicio que nos permite trascender los imperativos de la biología.

Creo que esto mismo es también verdad de nuestro uso de la razón teórica para determinar el carácter real del mundo que nos es presentado por la percepción sensible[64]. Como he dicho, el juicio de que se puede confiar en nuestros sentidos porque confiar en ellos contribuye a la eficacia es legítimo, pero el juicio de que se puede confiar en nuestra razón porque confiar en ella contribuye a la eficacia es incoherente. Ese juicio no puede él mismo depender de esta clase de confirmación empírica sin generar un retorno: emitir el juicio es necesariamente tomarlo como autoridad por su propio derecho. No creo que se disponga de ninguna otra postura mental en el ámbito teórico.

En el ámbito del valor y la razón práctica, creo que es coherente ser subjetivista, considerar todas las impresiones de valor objetivo o razones objetivas para la acción como ilusorias y pensar de los procesos de deliberación práctica y razonamiento moral como nada más que formas sofisticadas de decidir lo que uno realmente quiere. El solo uso teórico de la razón proporciona un marco dentro del cual es posible en principio pensar el sistema motivacional propio solo como un producto de causas naturales y nunca como una respuesta a lo que es realmente bueno o malo. Aunque encuentro imposible adoptar esta posición creo que es comprensible. El problema es de plausibilidad relativa. Y tengo que reconocer que para alguien que no está dispuesto a aceptar el realismo del valor las consecuencias radicales que he dibujado de él solamente incrementan su implausibilidad.

CONCLUSIÓN

La filosofía tiene que proceder comparativamente. Lo mejor que podemos hacer es desarrollar las concepciones alternativas rivales en cada uno de sus puntos importantes tan completa y cuidadosamente como sea posible, dependiendo de nuestras previas afinidades, y ver cómo dan la talla. Esa es una forma más creíble de progreso que hallar una prueba decisiva o que la refutación.

En el presente clima de un naturalismo científico dominante, fuertemente dependiente de las explicaciones especulativas darwinistas sobre prácticamente todo y armado hasta los dientes contra ataques de la religión, he pensado que sería útil especular sobre alternativas posibles. Sobre todo, me gustaría ampliar las fronteras de lo que no se considera como impensable a la luz de lo poco que realmente conocemos sobre el mundo. Habría un avance si el *establishment* teórico laico y la cultura ilustrada contemporánea que lo domina pudieran distanciarse de las lagunas del materialismo y del darwinismo, aplicarse sus propias etiquetas peyorativas. He intentado mostrar que este acercamiento es incapaz de proporcionar una explicación adecuada de nuestro universo, tanto constitutiva como históricamente.

Sin embargo, estoy seguro de que mi propio intento de explorar alternativas tiene mucha falta de imaginación. Una comprensión del universo como fundamentalmente propenso a engendrar vida y mente requerirá probablemente un punto de partida mucho más radical que las formas al uso de explicación naturalista que soy capaz de concebir de momento. Específicamente, al intentar comprender la conciencia como fenómeno biológico, es demasiado fácil olvidar la radical diferencia entre lo subjetivo y lo objetivo y caer en el error de pensar lo mental en términos

tomados de nuestras ideas de los eventos y procesos físicos. Wittgenstein fue consciente de este error aunque su forma de evitarlo a través de una exploración de la gramática del lenguaje mental me parece sencillamente insuficiente.

Es perfectamente posible que la verdad esté más allá de nuestro alcance, en virtud de nuestras limitaciones cognoscitivas intrínsecas y no solo más allá de nuestro entendimiento en el escenario actual de desarrollo intelectual de la humanidad. Pero creo que eso no podemos saberlo y que tiene sentido continuar buscando una comprensión sistemática de cómo nosotros y otros seres vivos cabemos en el mundo. En este proceso, la capacidad de generar y rechazar hipótesis falsas desempeña una función esencial. He argumentado pacientemente contra la forma prevalente de naturalismo, un materialismo reduccionista que pretende aprehender la vida y la mente mediante su extensión neodarwinista. Pero, volviendo a mis incisos introductorios, encuentro este punto de vista, de entrada, increíble, como un heroico triunfo de una teoría ideológica sobre el sentido común. La evidencia empírica se puede interpretar para que albergue diferentes teorías comprensivas, pero en este caso el coste en requiebros probabilistas y conceptuales es prohibitivo. Apostaría a que el actual consenso bien pensante será risible en una generación o dos, aunque por supuesto pueda ser sustituido por un nuevo consenso que sea tan poco válido como él. La voluntad humana de creer es inagotable.

ÍNDICE ANALÍTICO Y ONOMÁSTICO

A

B

C

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

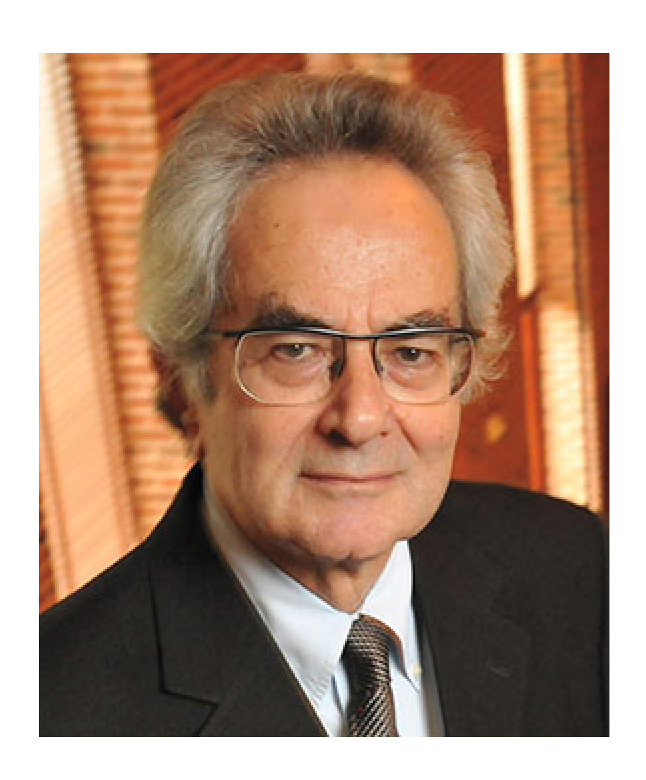
S

T

U

V

W



THOMAS NAGEL (Belgrado, 1937). Nacido en Belgrado, en la actual Serbia, de familia judía, se nacionalizó norteamericano en 1944. Estudió en bachillerato en Nueva York y filosofía en la Universidad de Oxford, doctorándose en la norteamericana de Harvard (1963), con una tesis dirigida por John Rawis. Comenzó la docencia universitaria en Berkeley (1963-66), de donde pasó a Pinceton (1966-80) y, finalmente, a la City University de Nueva York.

Está considerado uno de los mayores expertos mundiales en la filosofía de la mente. Es autor de *The Possibility of Altruism* (1970), *The View from Nowhere* (1986), *What Does It All Mean?: A Very Short Introduction to Philosophy* (1987), *Equality and Partiality* (1991), *Other Minds: Critical Essays*, 1969-1994 (1995), *The Last Word* (1997), *The Myth of Ownership: Taxes and Justice* (2002), *Concealment and Exposure and Other Essays* (2002).

En lengua española han sido editados, entre otros: *La muerte en cuestión. Ensayos sobre la vida humana* (México DF, FCE, 1981), *Igualdad y parcialidades. Bases éticas de la teoría política* (Barcelona, Paidós, 1996),

Una visión de ningún lugar (México DF, FCE, 1996), *¿Qué significa todo esto?* (Mexico DF, FCE, 1996), *Otras mentes. Ensayos críticos, 1969-1994* (Barcelona, Gedisa, 2000), *La última palabra* (Barcelona, Gedisa, 2001), *La posibilidad del altruismo* (México DF, FCE, 2004).

El pensamiento

Filósofo de la mente, indaga en los planos de la ética y la justicia universal y de aquellas expresiones que permanecen en la Humanidad, enfrentándose así a las visiones neodarwinistas y relativistas y, en general, al pensamiento postmoderno.

El análisis de la objetividad es uno de los principales focos de atención de Nagel. Hasta dónde es posible la objetividad y cuál es la relación entre conciencia, verdad y objetividad ocupan su reflexión. La objetividad es inelúdible en la búsqueda de la verdad científica, por ejemplo, y requiere un ejercicio de distanciamiento y anulación del plano subjetivo del investigador. Pero, en general, la subjetividad forma parte de la realidad, de las formas de percibir el mundo, por lo que la realidad fluctúa entre ambas perspectivas, entre ambos polos, señala en su libro *The View from Nowhere* (1986), siendo el "no lugar" la objetividad imposible, la visión sin el referente subjetivo de quien ve.

Analiza la relación mente-cerebro a partir de la existencia de la conciencia como hecho diferencial que, entiende, no puede ser reducido al análisis de la física, ya que excluye lo subjetivo, lo diferencial. La conciencia de ser proyecta la experiencia subjetiva de la experiencia en la percepción de la realidad. Podemos intentar objetivar cuál es la percepción que tiene un murciélago, explica en su célebre artículo "What is it like to be a bat?" (1974), pero nos falta, para ser plenamente objetivos, sentir la subjetividad perceptiva, la conciencia de ser del murciélago.

Notas

[1] Para una declaración de principios clara véase Steven Weinberg, *Dreams of a Final Theory* (Nueva York, Pantheon Books, 1992), capítulo 3. <<

[2] Véase Richard Dawkins, *The Blind Watchmaker. Why the Evidence of Evolution Reveals a Universe without Design* (Nueva York, Norton, 1986) para una exposición canónica que parece convencer prácticamente a todos. <<

[3] Para una esclarecedora cuenta de los propios puntos de vista de Darwin sobre las formas más básicas de explicación véase Elliot Sober, «Darwin's Discussions of God», en *Did Darwin Write the «Origin» Backwards? Philosophical Essays on Darwin's Theory* (Amherst, NY, Prometheus Books, 2011), 121-128. <<

[4] Véase Dawkins, *The Blind Watchmaker,* 77-86. Jerry Fodor y Massimo Piatelli-Palmarini argumentan en la primera parte de su libro *What Darwin Got Wrong* (Nueva York, Farrar, Straus & Giroux, 2010) que la teoría evolucionista darwiniana asigna demasiada carga de explicación de la naturaleza funcional de los organismos a la influencia externa de la selección natural y no la otorga suficientemente a las fuentes de la variación genética. Este punto es independiente de su ataque a la presunta intencionalidad de la idea de selección natural en la segunda parte de su libro, que me parece, como a otros, que se basa en un error de interpretación.

Hay también más figuras principales que insisten en que la evidencia clama por una explicación más restrictiva de las fuentes de variación en el material genético. Marc W. Kirschner y John C. Gerhart en *The Plausibility of Life: Resolving Darwin's Dilemma* (New Haven, CT, Yale University Press, 2005) sugieren que la variación genética está sesgada para facilitar el cambio evolutivo, aunque insisten en que esto no implica una corrección de una concepción reduccionista de la naturaleza más amplia. Stuart Kauffman sugiere en varios libros que la variación no se debe a la casualidad y que los principios de la auto-organización espontánea desempeñan una función más importante que la selección natural en la historia evolutiva. Véase *At Home in the Universe: The Search for Laws of Self-Organization and Complexity* (Nueva York, Oxford University Press, 1995); *Investigations* (Nueva York, Oxford University Press, 2000); *Reinventing the Sacred: A New View of Science, Reason and Religion* (Nueva York, Basic Books, 2008). <<

[5] En verdad puede haber algo profundamente confuso en la petición de una explicación tal, por una razón señalada por Roger White y que discuto en el capítulo 4. <<

[6] Véase Michael J. Behe, *Darwin’s Black Box: The Biochemical Challenge to Evolution* (Nueva York, Simon & Schuster, 1996); Behe, *The Edge of Evolution: The Search for the Limits of Darwinism* (Nueva York, Free Press, 2007); Stephen C. Meyer, *Signature in the Cell: DNA and the Evidence for Intelligent Design* (Nueva York, Harper One, 2009). <<

[7] Véase David Berlinski, «On the Origins of Life», *Commentary*, febrero de 2006, reimpreso en Berlinski, *The Deniable Darwin and Other Essays* (Seattle, Discovery Institute Press, 2009). Véase también Brian Goodwin, *How the Leopard Changed Its Spots: The Evolution of Complexity* (Nueva York, Scribner's, 1994). <<

[8] Hay también críticas de teorías actuales de algunos de los que a pesar de todo se espera una solución reductiva, por ejemplo Robert Shapiro, *Origins: A Skeptic’s Guide to the Creation of Life on Earth* (Nueva York, Summit Books, 1986); Shapiro «A Simpler Origin for Life», *Scientific American*, 12 de febrero de 2007. Una explicación muy clara de los múltiples aspectos de la investigación actual sobre el origen de la vida y la posibilidad de vida extraterreste se encuentra en Steven Benner, *Life, the Universe and the Scientific Method* (Gainesville, FL, FLAME Press, 2008). Aunque da por supuesto que es una tarea para la química dice (287): «Existe una posibilidad real de que la teoría actual no pueda *nunca* resolver el problema que nos ocupa dejando abierta la posibilidad para una verdadera revolución en las ciencias relacionadas y circundantes». Por supuesto, el no quiere decir diseño inteligente.

Un problema con la mayoría de la investigación actual más destacada es que la síntesis de los componentes individuales del material genético está tan fuertemente controlada y guiada por los experimentadores que proporciona escasa evidencia sobre qué procesos podrían haber ocurrido sin guía inteligente. Y el problema crucial de cómo podrían estos componentes haberse combinado en una secuencia codificada rica en información se ha dejado desatendido. <<

[9] La bibliografía es extensa. Véase por ejemplo, Kenneth R. Miller, *Finding Darwin's God: A Scientist's Search for Common Ground between God and Evolution* (Nueva York, Cliff Street, 1999); Philip Kitcher, *Living with Darwin: Evolution, Design, and the Future of Faith* (Nueva York, Oxford University Press, 2007); Elliot Sober, *Evidence and Evolution: The Logic Behind the Science* (Cambridge, UK, Cambridge University Press, 2008); para una muestra de los dos lados del debate ver Neil A. Manson (ed.), *God and Design: The Teleological Argument and Modern Science* (Nueva York, Routledge, 2003). <<

[10] No solo soy poco receptivo a esa idea sino, como he dicho en algún sitio más, profundamente adverso. <<

[11] Este programa ha sido desarrollado con dedicación en los escritos de Daniel Dennett. <<

[12] Aunque Einstein parece haberlo considerado como una cuestión abierta cuando señala: «¿Tuvo Dios elección cuando creó el universo?». <<

[13] Véase Barry Stroud *The Significance of Philosophical Scepticism* (Oxford, Clarendon, 1984), cap. 6, «Naturalized Epistemology». <<

[14] Alvin Plantinga, *Warrant and Proper Function* (Nueva York, Oxford University Press, 1993), cap. 12. <<

[15] Sharon Street, «A Darwinian Dilemma for Realist Theories of Value», *Philosophical Studies*, 127, núm. 1 (enero de 2006), 209-266. <<

[16] Véase P. F. Strawson, *Individuals: An Essay in Descriptive Metaphysics* (Londres, Methuen, 1959). <<

[17] Veo con muy buenos ojos la siguiente declaración de Jaegwon Kim: «La metafísica es el ámbito en el que se aúnan diferentes lenguajes, teorías, explicaciones y sistemas conceptuales y ponen en orden y clarifican sus mutuas relaciones ontológicas. Que hay tal ámbito común es la creencia de un realismo amplio y no tendencioso sobre nuestras actividades cognoscitivas. Si no crees que haya tal ámbito, bien, eso es también metafísica». *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* (Cambridge, MA, MIT Press, 1998), 66. <<

[18] Gilbert Ryle, *The Concept of Mind* (Londres, Hutchinson, 1949); Ludwig Wittgenstein, *Philosophical Investigations*, trad. G. E. M. Anscombe (Nueva York, Macmillan, 1953). <<

[19] U. T. Place, «Is Consciousness a Brain Process?», *British Journal of Psychology*, 47 (1956), 44-50; J. J. C. Smart, «Sensations and Brain Processes», *Philosophical Review*, 68 (1959), 141-156. <<

[20] Esta es la objeción de Max Black a la teoría de la identidad, citada en Smart, «Sensations and Brain Processes», pág. 148, nota 11. <<

[21] Véase D. M. Amstrong, *A Materialist Theory of the Mind* (Londres, Routledge & Kegan Paul, 1968). <<

[22] Saul A. Kripke, *Naming and Necessity* (Cambridge, MA, Harvard University Press, 1980). <<

[23] René Descartes, *Meditación*, 6. <<

[24] Véase Paul K. Feyerabend, «Mental Events and the Brain», *Journal of Philosophy*, 60 (1963), 295-296. <<

[25] He dicho más sobre esto en «The Psychophysical Nexus» en *Concealment and Exposure, and Other Essays* (Nueva York, Oxford University Press, 2002), 194-235. <<

[26] Los biólogos evolutivos hacen ahora hincapié en que las fuentes de variación no son «aleatorias» en el sentido ordinario, que las mutaciones en el material genético sobre el cual actúa la selección natural están restringidas de varias formas y no son igualmente probables y que esto contribuye significativamente al proceso. Por ejemplo, los extraordinarios descubrimientos recientes de la biología evolutiva del desarrollo *(evo-devo)* parecen implicar que hay mucho más sistema y menos azar en las fuentes de la variación genética de lo que se había supuesto con anterioridad. Pero tales hechos tendrían también que explicarse en último lugar por principios físicos en una teoría reduccionista. <<

[27] A día de hoy, muchos investigadores creen que un «mundo de ARN» precedió a la aparición del ADN, pero pueden aplicarse los mismos problemas de probabilidad. <<

[28] Cfr. la reseña de Richard Lewontin del libro de Carl Sagan *The Demon-Haunted World*, en *New York Review of Books*, 9 de enero de 1997: «No es que los métodos e instituciones de la ciencia nos obliguen de alguna manera a aceptar una explicación material del mundo fenoménico, sino que, por el contrario, estamos forzados por nuestra adherencia *a priori* a las causas materiales a crear un aparato de investigación y un conjunto de conceptos que produzcan explicaciones materiales, no importa qué contra-intuitivo pueda ser, no importa cómo de místico resulte a los no iniciados. Es más, ese materialismo es absoluto porque no podemos permitirnos que Dios ponga un pie en la puerta». <<

[29] Pero el dualismo de la sustancia dejaría a la biología con un enorme problema similar al que estamos discutiendo ahora: a saber, ¿por qué ha producido la evolución física organismos de un cierto tipo capaces de ser ocupados por e interactuar con las mentes? Presumiblemente la respuesta dependería de asignar a la substancia mental una función en la adaptación de los organismos sucesivos. Pero, desde un punto de vista evolucionista, permanecería siendo un completo misterio por qué aparecen mentes del tipo apropiado y se unen a organismos. <<

[30] Stephen Jay Gould y Richard C. Lewontin, «The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme», *Proceedings of the Royal Society of London*, B 205, 581-598. <<

[31] Aquí hay un peligro de confusión terminológica. Usaré «*reductivo*» como el término general para las teorías que analizan las propiedades de todos complejos en las propiedades de sus elementos más básicos. Continuaré usando «*reduccionista*» para el tipo más específico de teorías reductivas que analizan fenómenos de un nivel más alto exclusivamente en términos de elementos físicos y de sus propiedades físicas. El reduccionismo psicofísico es un ejemplo. El punto que hay que recordar es que es posible que una teoría antirreduccionista sea reductiva dado que los elementos a los que reduce los fenómenos de más alto nivel no son exclusivamente físicos. Esta es la clase de teoría reductiva de la que estoy hablando aquí. <<

[32] Tom Sorell, *Descartes Reinvented* (Cambridge, UK, Cambridge University Press, 2005), 95. <<

[33] Para una detallada y esclarecedora discusión sobre este tema, véase Galen Strawson et al., *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?* (Exeter, UK, Imprint Academic, 2006). El libro consiste en un ensayo central de Strawson, diecisiete comentarios de otros sobre él y una larga réplica de Strawson elaborando su defensa del pansiquismo. Véase También mis ensayos «Panpsychism», en *Mortal Questions* (Cambridge, UK, Cambridge University Press, 1979) y «The Phychophysical Nexus»; y la aguda e históricamente documentada argumentación de Charles Hartshorne, «Physics and Psychics: The Place of Mind in Nature», en *Mind in Nature: Essays on the Interface of Science and Philosophy*, John B. Cobb, Jr. y David Ray Griffin (eds.) (Washington, DC, University Press of America, 1977), 89-96.

La filosofía del organismo de Alfred North Whitehead pertenece también a esta línea. Whitehead argüía que identificar las abstracciones de la física con toda la realidad era cometer la falacia de la concreción fuera de lugar y que las entidades concretas, todas hasta el nivel de los electrones, deberían entenderse como incorporando de alguna manera un punto de vista sobre el mundo. Véase Whitehead, *Science and the Modern World: Lowell Lectures, 1925* (Nueva York, Macmillan, 1925). <<

[34] Estoy en deuda aquí con Sharon Street. <<

[35] Para una mayor exposición, véase Nagel «The Psichophysical Nexus». <<

[36] Aunque, por supuesto, Aristóteles no tuvo nuestra concepción de una evolución histórica del mundo a lo largo del tiempo. <<

[37] Véase John Hawthorne y Daniel Nolan, «What Would Teleological Causation Be?», en *Metaphysical Essays*, John Hawthorne (ed.) (Oxford, Oxford University Press, 2006) 265-283. <<

[38] John Leslie defiende la posibilidad de que el valor pueda explicar la existencia Véase «The Theory That the World Exists because it Should», *American Philosophical Quarterly*, 7 (1970), 286-298, e *Infinite Minds: A Philosophical Cosmology* (Oxford, Clarendon, 2001). <<

[39] Véase Sharon Street, «A Darwinian Dilemma for Realist Theories of Value», *Philosophical Studies*, 127 (2006), 109-166. Desarrollo este asunto en el próximo capítulo. <<

[40] Véase Elliott Sober y David Sloan Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA, Harvard University Press, 1998). <<

[41] Véase el último capítulo de mi *The Last Word* (Nueva York, Oxford University Press, 1997); aquí estoy prosiguiendo la discusión de los problemas planteados allí. <<

[42] Para referencias relevantes véase nota 8 del capítulo 1 y también Iris Fry, *The Emergence of Life on Earth: A Historical and Scientific Overview* (New Brunswick, NJ, Rutgers University Press, 2000). <<

[43] Roger White, «Does Origins of Life Research Rest on a Mistake?», *Noûs*, 41 (2003), 475. <<

[44] Ibíd., 467. <<

[45] Véase Hawthorne and Nolan, «What Would Teleological Causation Be?». <<

[46] Véase Alvin Plantinga, *Where the Conflict Really Lies: Science, Religion, and Naturalism* (Nueva York, Oxford University Press, 2011). <<

[47] ¿Hay otras alternativas? Bien, está la hipótesis de que este universo no es único, sino de que existen todos los universos posibles y nos encontramos, no sorprendentemente, en uno que contiene vida. Pero esto es un escurrir el bulto que dispensa del intento de explicar nada. Y sin la hipótesis de los universos múltiples, la observación de que si la vida no hubiera llegado a existir nosotros no estaríamos aquí no es significativa. Uno no muestra que algo no requiere explicación señalando que es una condición de la existencia de uno. Si pido una explicación del hecho de que el aire está presurizado al nivel del mar en un avión transcontinental, no es respuesta señalar que si no lo estuviera estaríamos muertos. Cfr. Leslie, *Infinite Minds*, 207. Véase también Roger White, «Fine-Tuning and Multiple Universes», *Noûs*, 34 (2000), 260-276; y Ian Hacking, «The Inverse Gambler's Fallacy: The Argument from Design. The Anthropic Principle Applied to Wheeler Universes», *Mind*, 96 (1987), 331-340 para ulteriores problemas con la apelación a los múltiples universos para explicar el carácter de este. <<

[48] Dejo a un lado el amplio tema de las verdades normativas contingentes, sobre las que tanto los realistas como los subjetivistas acerca del valor pueden estar de acuerdo en que son dependientes de nuestras respuestas colectivas. Cualquiera que ha vivido lo bastante para ver cómo el lenguaje se diluye sabrá lo que quiero decir. *Disinterested* viene a significar *uninterested*, *precipitous* viene a significar *precipitate, criteria* se convierte en singular y *I can't help but think* y *I could careless* son ubicuos. Leyendo con exasperación diaria publicaciones canónicas como el *New York Times*, tengo que reconocer mi inconmovible convicción de que estos son errores producto de un conjunto contingente de normas lingüísticas internalizadas hace años de las que no puedo librarme (y no lo haría incluso si pudiera ya que son, por supuesto, correctas). <<

[49] Estoy en deuda en este punto con Matthew Silverstein. <<

[50] Véase T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA, Harvard University Press, 1998), 56-57. <<

[51] Otra forma de establecer la respuesta del realista es decir que la pregunta incorpora un error categorial, en el sentido de Ryle. <<

[52] A lo largo de la argumentación continuaré hablando en términos muy generales de la explicación darwinista, a pesar de que la teoría evolucionista darwinista es inmensamente compleja y objeto de importantes controversias internas. Creo que para los propósitos presentes las complejidades y desacuerdos dentro de la teoría darwinista no importan. También, como en el capítulo previo, dejaré a un lado todas las dudas sobre la verdad de la teoría evolucionista y supondré que hay una versión suya que pueda dar cabida a la conciencia. <<

[53] *Philosophical Studies*, 127 (2006), 109-166. <<

[54] Cfr. Street, «A Darwinian Dilemma», 123-125 y 135. <<

[55] Véase Sharon Street, «Constructivism about Reasons», *Oxford Studies in Metaethics*, 3 (2008), 207-246. <<

[56] Scanlon, *What We Owe to Each Other*, 20. <<

[57] Para una explicación persuasiva de la relación entre conciencia y voluntad libre, véase David Hodgson, *Rationality + Consciousness = Free Will* (Nueva York, Oxford University Press, 2011). <<

[58] Scanlon, *What We Owe to Each Other*, 68. <<

[59] «Imaginemos, por ejemplo, que hemos evolucionado más en la línea de los leones, de manera que los machos en circunstancias relativamente frecuentes tienen una fuerte tendencia valorativa irreflexiva a aniquilar la descendencia que no es suya como ‘demandados’ por las circunstancias y las hembras, a su vez, no experimentan una fuerte tendencia irreflexiva a ‘ir contra’ un macho cuando asesina a su descendencia en tales circunstancias, por el contrario llega a ser receptiva a sus requerimientos muy pronto». <<

[60] Véase la descripción de la lógica de tal explicación en el capítulo previo. <<

[61] Tal posibilidad es descrita por C. D. Broad en *The Mind and Its Place in Nature* (Londres, Routledge & Kegan Paul, 1925), 81-94. La concepción de Henri Bergson de la evolución creadora postuló una tendencia similar, pero él la pensó no como una adición al orden natural sino como la libre creación de una fuerza vital universal. Véase *L'évolution créatrice* (1907); disponible en inglés como *Creative Evolution*, trad. Arthur Mitchell (Nueva York, Henry Holt, 1911). [*N. del T.* Existe traducción española, *La evolución creadora*, Madrid, Espasa-Calpe, 1973. Trad. María Luisa Pérez Torres.] <<

[62] Cfr. la explicación de Derek Parfit de por qué existe el universo. «Why Anything? Why This?», *London Review of Books*, 22 de enero de 1992; reimpreso en Parfit, *On What Matters* (Oxford, Oxford University Press, 2011). <<

[63] Francis Crick, *Life Itself: Its origin and Nature* (Nueva York, Simon & Schuster, 1981), 88. Véase también Jacques Monod, *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*, trad. Austin Wainhouse (Nueva York, Knopf, 1971) [*N. del T*. Disponible en castellano como *El azar y la necesidad: ensayo sobre la filosofía natural de la biología moderna*, Barcelona, Tusquets, 1981. Trad. Francisco Ferrer Lerín] y Ernst Mayr, *Evolution and Diversity of Life: Selected Essays* (Cambridge, MA, Harvard University Press, 1976). <<

[64] Sharon Street, de una forma muy interesante, sostiene que la razón teórica tampoco puede interpretarse realistamente si nuestras capacidades de razonamiento tienen un alcance significativo en una explicación evolucionista. Ella es realista sobre la *verdad* respecto del mundo natural, pero no es realista sobre las razones epistémicas. En el ámbito del valor, sin duda, es antirrealista tanto sobre la verdad como sobre las razones. Véase Sharon Street, «Evolution and the Normativity of Epistemic Reasons», *Canadian Journal of Philosophy*, Supplementary Volume 35 (2011), 213-248. <<